

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

```

AC AAB12236;
XX
XX 10-NOV-2000 (first entry)
DE Partial sequence of HIV-1 strain ESS gp41 immunodominant region.
XX
XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;
KM acquired immunodeficiency syndrome; group O HIV; gp41; ESS.
XX
XX Human immunodeficiency virus type 1.
OS
XX EP1013766-A2.
PN
XX 28-JUN-2000.
PD
XX
XX 29-NOV-1999; 99EP-0309491.
XX
XX 30-NOV-1998; 98US-0110292.
PR 08-FEB-1999; 99US-0119138.
PR 04-NOV-1999; 99US-0433428.
XX
XX (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.
XX
XX De Lays R, Zheng J;
XX
XX WPI; 2000-402205/35.
DR
XX
XX New antigenic peptides and peptide functional derivatives, useful for
PT detection of antibodies produced in response to human immunodeficiency
PT virus group O antibodies -
XX
XX Example 1; Fig 1; 52pp; English.
PS
XX
XX The present sequence is a partial gp41 protein of human immunodeficiency
CC virus type 1 (HIV-1) strain ESS. HIV is the principle aetiological
CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV
CC envelope protein, and may be used as an antigen for the detection of
CC antibodies produced in response to HIV infection. ESS is a member of
CC HIV group O (outlier). The present sequence is the immunodominant region
CC of gp41. This sequence was used in a sequence homology alignment,
CC which in turn was used to derive a consensus sequence peptide: peptide
CC 147 (AAB12235).
XX
XX
XX Sequence 33 AA:
SQ
XX
XX Query Match 75.7%; Score 131; DB 21; Length 33;
XX Best Local Similarity 75.0%; Pred. No. 8.2e-12;
XX Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
XX
XX 3 ETLMDQDRLNSMCGCKGRICYSARWH 30
QY |||:|||||:|||||:|||||:|
Db 6 etllqngqllnswgckgrivcytsvkw 33
XX
XX
XX RESULT 12
XX AAM07343
XX ID AAM07343 standard; peptide: 40 AA.
XX
XX AAM07343;
XX
XX 03-JUN-1997 (first entry)
XX
XX Partial sequence of gp41 from HIV-1 gp. O strain BCF02 (ESS).
DE
XX
XX Human immunodeficiency virus type 1; HIV-1; envelop; group M; group O;
KM C2V3-env; gp41; gag; retrovirus; strain; gp120; hypervariable loop;
XX primer; hybridisation; amplification; PCR; polymerase chain reaction;
XX immunogen; antibody.
XX
XX Human immunodeficiency virus type 1.
XX
XX OS
XX *XX W09627013-
XX PN

```

```

XX
XX 06-SEP-1996.
XX
XX 26-FEB-1996; 96WO-FR00294.
PD
XX
XX 27-FEB-1995; 95FR-0002236.
PR
XX
XX (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Chaix-Baudier ML, Lousset-Ajaka I, Ly T, Saragosti S, Simon F;
PI WPI; 1996-412779/41.
XX
XX N-PSDB; AAT44918.
DR
XX
XX New strains of HIV-1 group O, related DNA fragments, peptide(s) and
PT antibodies - useful for diagnosis, screening and typing, or as
PT immunogens
XX
XX Claim 12; Page 33; 71pp; French.
PS
XX
XX Human immunodeficiency virus type 1 (HIV-1) strains are currently
CC divided into 2 major groups based on the nucleotide sequences of the
CC envelop gene (env): group M containing sub-groups A-G, and group O
CC containing the strains AM70 and WPS180. The invention relates to the
CC discovery of several new strains of HIV-1 which can be placed in group O,
CC based on the partial sequences of the C2V3-env, gp41 and gag genes (see
CC AM744907-39 and AAM07329-64). The novel strains have been deposited as
CC retroviruses CCMC I-1544 (BCF02 (ESS)), 1543 (BCF01 (FAN)), 1546 (BCF07
CC (MAN)), 1547 (BCF08 (NKO)) and 1545 (BCF03 (POC)). The sequence
CC presented here is from the strain BCF02 (ESS) and corresponds to a
CC fragment of the gp41 protein encoded by the env gene. The nucleic acids
CC can be used to detect gp. O HIV-1 strains by hybridisation or (as
CC primers) by gene amplification, also for screening and typing of such
CC strains. Peptides encoded by the nucleic acids can be used as immunogens
CC to raise Ab for detecting gp. O HIV-1.
XX
XX
XX Sequence 40 AA:
SQ
XX
XX Query Match 75.7%; Score 131; DB 17; Length 40;
XX Best Local Similarity 75.0%; Pred. No. 1e-11;
XX Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
XX
XX 3 ETLMDQDRLNSMCGCKGRICYSARWH 30
QY |||:|||||:|||||:|||||:|
Db 11 etllqngqllnswgckgrivcytsvkw 38
XX
XX
XX RESULT 13
XX AAM80469
XX ID AAM80469 standard; peptide: 32 AA.
XX
XX AAM80469;
XX
XX 28-JAN-1999 (first entry)
XX
XX Peptide derived from a conserved sequence of group O human HIV.
DE
XX
XX Group O human immune deficiency virus; HIV; detection; infection.
XX
XX Synthetic.
XX OS Immune deficiency virus.
XX
XX W09845323-A1.
PN
XX
XX 15-OCT-1998.
PD
XX
XX 06-APR-1998; 98WO-FR00691.
XX
XX 24-FEB-1998; 98FR-0002212.
PR 09-APR-1997; 97FR-0004356.
XX
XX
XX
XX

```

CC a HIV envelope protein, and so the present sequence may be used as an  
CC antigen for the detection of antibodies produced in response to HIV  
infection.

XX Sequence 23 AA;

#### Query Match

Best Local Similarity 77.5%; Score 134; DB 21; Length 23;  
Pred. No. 2,1e-12;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 DOORLNSWCGKGRITCYTSARWH 30  
:|||||  
Db 1 egqrlnswgckgrilicytsarwh 23

#### RESULT 9

AAB12254  
ID AAB12254 standard; peptide: 23 AA.

XX AAB12254;

DT 10-NOV-2000 (first entry)

DE HIV-1 gp41 immunodominant region consensus sequence peptide 147 # 1.

XX HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;  
KM acquired immunodeficiency syndrome; group O HIV; gp41.

XX Human immunodeficiency virus type 1.

OS EPI013766-A2.

XX 28-JUN-2000.

PF 29-NOV-1999; 99EP-0309491.

PR 30-NOV-1998; 98US-0110292.

PR 08-FEB-1999; 99US-0119138.

PR 04-NOV-1999; 99US-0433428.

PA (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.

PI De Leys R, Zheng J;

DR WPI: 2000-402205/35.

XX New antigenic peptides and peptide functional derivatives, useful for  
PT detection of antibodies produced in response to human immunodeficiency  
virus group O antibodies -

PS Claim 1; Page 35; 52pp; English.

XX The present sequence is peptide 147 of Human Immunodeficiency Virus Type  
CC 1 (HIV-1). This sequence is a partial consensus sequence of the  
CC immunodominant region of gp41 protein derived from a variety of HIV-1  
CC group O (outlier) strains: ANT70, WVP5180, VAV, DUR, POC, FAN, LOB, MAN,  
CC NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686, ABT063, ABT124, ABT1123,  
CC ABT2156, 193Ha, CDC7755, CDC1897, HLD28, 1515, 1516, D47-2d, HCYT2c,  
CC NT42 and PE41 (see AAB12207 to AAB12236). HIV is the principle  
CC aetiological agent for acquired immunodeficiency syndrome (AIDS). gp41 is  
CC a HIV envelope protein, and so the present sequence may be used as an  
CC antigen for the detection of antibodies produced in response to HIV  
infection.

XX Sequence 23 AA;

#### Query Match

Best Local Similarity 76.9%; Score 133; DB 21; Length 23;  
Pred. No. 2,9e-12;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 DOORLNSWCGKGRITCYTSARWH 30

Db :|||||  
1 egqrlnswgckgrilicytsarwh 23

#### RESULT 10

AAB12255  
ID AAB12255 standard; peptide: 23 AA.

XX AAB12255;

DT 10-NOV-2000 (first entry)

DE HIV-1 gp41 immunodominant region consensus sequence peptide 147 # 2.

XX HIV-1; AIDS; human immunodeficiency virus type 1; antibody detection;  
KM acquired immunodeficiency syndrome; group O HIV; gp41.

XX Human immunodeficiency virus type 1.

OS Human immunodeficiency virus type 1.

FT Key Location/Qualifiers

FT MISC-difference 1

PN EPI013766-A2.

PD 28-JUN-2000.

PF 29-NOV-1999; 99EP-0309491.

PR 30-NOV-1998; 98US-0110292.

PR 08-FEB-1999; 99US-0119138.

PR 04-NOV-1999; 99US-0433428.

PA (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.

PI De Leys R, Zheng J;

DR WPI: 2000-402205/35.

XX New antigenic peptides and peptide functional derivatives, useful for  
PT detection of antibodies produced in response to human immunodeficiency  
virus group O antibodies -

PS Claim 1; Page 35; 52pp; English.

XX The present sequence is a peptide 147 related peptide from Human  
CC Immunodeficiency Virus Type 1 (HIV-1). This sequence is a partial  
CC consensus sequence of the immunodominant region of gp41 protein derived  
CC from a variety of HIV-1 group O (outlier) strains: ANT70, WVP5180, VAV,  
CC DUR, POC, FAN, LOB, MAN, NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686,  
CC ABT063, ABT124, ABT1123, ABT2156, 193Ha, CDC7755, CDC1897, HLD28, 1515,  
CC 1516, D47-2d, HCYT2c, NT42 and PE41 (see AAB12207 to AAB12236). HIV is  
CC the principle aetiological agent for acquired immunodeficiency syndrome  
CC (AIDS). gp41 is a HIV envelope protein, and so the present sequence may  
CC be used as an antigen for the detection of antibodies produced in  
response to HIV infection.

XX Sequence 23 AA;

#### Query Match

Best Local Similarity 76.3%; Score 132; DB 21; Length 23;  
Pred. No. 4e-12;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 DOORLNSWCGKGRITCYTSARWH 30  
:|||||  
Db 2 qqrlnswgckgrilicytsarwh 23

#### RESULT 11

AAB12236  
ID AAB12236 standard; peptide: 33 AA.



RESULT 4  
 AAB12261 standard; peptide: 220 AA.  
 ID AAB12261  
 AC AAB12261;  
 DT 10-NOV-2000 (first entry)  
 DE HIV group M/ group O mosaic protein # 1.  
 XX  
 XX  
 XX HIV-1; AIDS; human immunodeficiency virus type 1;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41.  
 XX  
 OS Human immunodeficiency virus type 1.  
 PA EP1013766-A2.  
 PN 28-JUN-2000.  
 PD 29-NOV-1999; 99EP-0309491.  
 PF 30-NOV-1998; 98US-0110292.  
 PR 08-FEB-1999; 99US-0119138.  
 PR 04-NOV-1999; 99US-0433428.  
 XX  
 XX (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.  
 XX  
 XX De Leys R, Zheng J;  
 PI WPI: 2000-402205/35.  
 DR  
 XX  
 XX New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies -  
 CC  
 XX  
 PS Claim 15; Page 37-38; 52pp; English.  
 XX  
 CC Human Immunodeficiency Virus (HIV) is the principle aetiological  
 CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV  
 CC envelope protein, and may be used as an antigen for the detection of  
 CC antibodies produced in response to HIV infection. Mosaic gp41 proteins  
 CC were constructed, in which the immunodominant region of group M HIV was  
 CC replaced by the corresponding region from group O HIV. The mosaic gp41  
 CC proteins would be useful as antigens, used in the detection of anti-group  
 CC O HIV antibodies produced in response to HIV infection. The present  
 CC sequence is one such mosaic protein.  
 XX  
 SQ Sequence 220 AA;  
 Query Match 90.8%; Score 157; DB 21; Length 220;  
 Best Local Similarity 96.4%; Pred. NO. 1.1e-14;  
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 ETLMDQRLNSGCKGRITCTSRWH 30  
 DB 44 eELmqnqgrlnswgckgrilcytsarwh 71  
 RESULT 5  
 AAB12253 standard; peptide: 368 AA.  
 ID AAB12253  
 AC AAB12253;  
 DT 10-NOV-2000 (first entry)  
 DE DHFR-hES-MH fusion protein.  
 XX  
 XX HIV-1; AIDS; human immunodeficiency virus type 1; DHFR-hES-MH;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41.  
 XX

OS Human immunodeficiency virus type 1.  
 OS Unidentified.  
 PA EP1013766-A2.  
 PN 28-JUN-2000.  
 PD 29-NOV-1999; 99EP-0309491.  
 PF 30-NOV-1998; 98US-0110292.  
 PR 08-FEB-1999; 99US-0119138.  
 PR 04-NOV-1999; 99US-0433428.  
 XX  
 XX (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.  
 XX  
 XX De Leys R, Zheng J;  
 PI WPI: 2000-402205/35.  
 DR  
 XX  
 XX New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies -  
 CC  
 XX  
 PS Example 5; Fig 6; 52pp; English.  
 XX  
 CC Human Immunodeficiency Virus (HIV) is the principle aetiological  
 CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV  
 CC envelope protein, and may be used as an antigen for the detection of  
 CC antibodies produced in response to HIV infection. Mosaic gp41 proteins  
 CC were constructed, in which the immunodominant region of group M HIV was  
 CC replaced by the corresponding region from group O HIV. The mosaic gp41  
 CC proteins would be useful as antigens, used in the detection of anti-group  
 CC O HIV antibodies produced in response to HIV infection. The present  
 CC sequence is a dihydrofolate reductase (DHFR) fusion protein of one such  
 CC mosaic protein.  
 XX  
 SQ Sequence 368 AA;  
 Query Match 90.8%; Score 157; DB 21; Length 368;  
 Best Local Similarity 96.4%; Pred. NO. 2e-14;  
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 ETLMDQRLNSGCKGRITCTSRWH 30  
 DB 240 eELmqnqgrlnswgckgrilcytsarwh 267  
 RESULT 6  
 AAB12252 standard; peptide: 439 AA.  
 ID AAB12252  
 AC AAB12252;  
 DT 10-NOV-2000 (first entry)  
 DE DHFR-hENV-MH fusion protein.  
 XX  
 XX HIV-1; AIDS; human immunodeficiency virus type 1; DHFR-hENV-MH;  
 KW acquired immunodeficiency syndrome; group O HIV; gp41.  
 XX  
 OS Human immunodeficiency virus type 1.  
 OS Unidentified.  
 PA EP1013766-A2.  
 PN 28-JUN-2000.  
 PD 29-NOV-1999; 99EP-0309491.  
 PF 30-NOV-1998; 98US-0110292.  
 PR 08-FEB-1999; 99US-0119138.  
 PR 04-NOV-1999; 99US-0433428.

XX PS Claim 1; Page 36; 52pp; English.

CC CC The present sequence is peptide 147-5 from Human Immunodeficiency Virus Type 1 (HIV-1). This sequence is a partial consensus sequence of the

CC CC immunodominant region of gp41 protein derived from a variety of HIV-1

CC CC group O (outlier) strains: AN770, MWP5180, VAV, DOR, POC, FAN, LOB, MAN, NAN, ESS, NKO, BCF09, BCF12, BCF13, BCF14, 686, ABT063, ABT124, ABT1123, CC ABT2156, 193H, CDC7755, CDC1897, HLD28, 1515, 1516, D47-2d, HCV72c, CC N42 and PE41 (see AAB12207 to AAB12336). HIV is the principle

CC CC aetiological agent for acquired immunodeficiency syndrome (AIDS). gp41 is CC a HIV envelope protein, and so the present sequence may be used as an CC antigen for the detection of antibodies produced in response to HIV CC infection.

XX SQ .Sequence 30 AA:

Query Match 100.0%; Score 173; DB 21; Length 30;  
Best Local Similarity 100.0%; Pred. No. 6.9e-18;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRETLMODQRLNSWCGKGRITCYTSARWH 30  
|||||  
Db 1 gretlmqdgqrlnswcgkgrilicytsarwh 30

RESULT 2

AAB12259 ID AAB12259 standard; peptide: 35 AA.

XX AC AAB12259;

DT 10-NOV-2000 (first entry)

XX DE Group O HIV-1 gp41 replacement peptide # 1.

XX KW HIV-1: AIDS: human immunodeficiency virus type 1; antibody detection;

XX KM acquired immunodeficiency syndrome; group O HIV; gp41.

XX OS Human immunodeficiency virus type 1.

XX PA EPI013766-A2.

XX PN 28-JUN-2000.

XX PD 29-NOV-1999; 99EP-0309491.

XX PF 30-NOV-1998; 98US-0110292.

XX PR 08-FEB-1999; 99US-0119138.

XX PR 04-NOV-1999; 99US-0433428.

XX PS (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.

XX PI De Leys R, Zheng J;

XX PT WPI; 2000-402205/35.

XX DR New antigenic peptides and peptide functional derivatives, useful for

XX PT detection of antibodies produced in response to human immunodeficiency

XX PT virus group O antibodies -

XX PS Claim 1; Page 37; 52pp; English.

XX PS The present sequence is a group O Human Immunodeficiency Virus Type 1

CC CC (HIV-1) gp41 peptide. HIV is the principle aetiological agent for

CC CC acquired immunodeficiency syndrome (AIDS). gp41 is a HIV envelope

CC CC protein. The present sequence was derived from a sequence homologous

CC CC alignment of group O and group M HIV gp41 descending helix sequences.

CC CC The present sequence was used as a group O replacement peptide, where it

CC CC was used to construct mosaic gp41 proteins, in which the group M

CC CC immunodominant region was replaced by the present sequence. The mosaic

CC CC gp41 proteins (AAB12261 and AAB12262) would be useful as antigens, which

CC CC would be used for the detection of anti-group O HIV antibodies produced

CC CC in response to HIV infection.

XX SQ .Sequence 35 AA:

Query Match 90.8%; Score 157; DB 21; Length 35;  
Best Local Similarity 96.4%; Pred. No. 1.6e-15;  
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 ETLMQDQRLNSWCGKGRITCYTSARWH 30  
|||||  
Db 8 etlmqngqrlnswcgkgrilicytsarwh 35

RESULT 3

AAB12262 ID AAB12262 standard; peptide: 149 AA.

XX AC AAB12262;

DT 10-NOV-2000 (first entry)

XX DE HIV group M/ group O mosaic protein # 2.

XX KW HIV-1: AIDS: human immunodeficiency virus type 1;

XX KM acquired immunodeficiency syndrome; group O HIV; gp41.

XX OS Human immunodeficiency virus type 1.

XX PN EPI013766-A2.

XX PD 28-JUN-2000.

XX PF 29-NOV-1999; 99EP-0309491.

XX PR 30-NOV-1998; 98US-0110292.

XX PR 08-FEB-1999; 99US-0119138.

XX PR 04-NOV-1999; 99US-0433428.

XX PS (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.

XX PI De Leys R, Zheng J;

XX PT WPI; 2000-402205/35.

XX DR New antigenic peptides and peptide functional derivatives, useful for

XX PT detection of antibodies produced in response to human immunodeficiency

XX PT virus group O antibodies -

XX PS Claim 15; Page 38-39; 52pp; English.

XX PS Human Immunodeficiency Virus (HIV) is the principle aetiological

CC CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV

CC CC envelope protein, and may be used as an antigen for the detection of

CC CC antibodies produced in response to HIV infection. Mosaic gp41 proteins

CC CC were constructed, in which the immunodominant region of group M HIV was

CC CC replaced by the corresponding region from group O HIV. The mosaic gp41

CC CC proteins would be useful as antigens, used in the detection of anti-group

CC CC O HIV antibodies produced in response to HIV infection. The present

CC CC sequence is one such mosaic protein.

XX SQ .Sequence 149 AA:

Query Match 90.8%; Score 157; DB 21; Length 149;  
Best Local Similarity 96.4%; Pred. No. 7.5e-15;  
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 ETLMQDQRLNSWCGKGRITCYTSARWH 30  
|||||  
Db 44 etlmqngqrlnswcgkgrilicytsarwh 71

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 15:26:17 ; Search time 29.93 Seconds  
(without alignments)  
111.334 Million cell updates/sec

Title: US-09-605-573a-69

Perfect score: 173  
Sequence: 1 GRETLMQDOORLNSWCKGRICYSARWH 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.032802:\*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

\* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	173	100.0	30	21	AA12264	HIV-1 gp41 immunod
2	157	90.8	35	21	AA12259	Group O HIV-1 gp41
3	157	90.8	149	21	AA12262	HIV group M/ group
4	157	90.8	220	21	AA12261	HIV group M/ group
5	157	90.8	368	21	AA12253	DHER-hes-MH fusio
6	157	90.8	439	21	AA12252	DHER-hes-MH fusio
7	155	89.6	28	21	AA12257	HIV-1 gp41 immunod
8	134	77.5	23	21	AA12256	HIV-1 gp41 immunod
9	133	76.9	23	21	AA12254	HIV-1 gp41 immunod
10	132	76.3	23	21	AA12255	HIV-1 gp41 immunod
11	131	75.7	33	21	AA12236	Partial sequence o

12	131	75.7	40	17	AA07343	Partial sequence o
13	130	75.1	32	19	AA08046	Peptide derived fr
14	130	75.1	33	21	AA12229	Partial sequence o
15	130	75.1	36	21	AA167606	Peptide #6 for det
16	130	75.1	113	20	AA105554	HIV-1 group O isol
17	129	74.6	33	21	AA12231	Partial sequence o
18	129	74.6	40	17	AA07346	Partial sequence o
19	128	74.0	33	21	AA12208	Partial sequence o
20	128	74.0	33	21	AA12215	Partial sequence o
21	128	74.0	33	21	AA12216	Partial sequence o
22	128	74.0	33	21	AA12218	Partial sequence o
23	128	74.0	33	21	AA12219	Partial sequence o
24	128	74.0	35	15	AA161690	HIV-type virus WVP
25	128	74.0	35	20	AA16167	HIV isolate WVP-51
26	128	74.0	36	20	AA16167	HIV isolate WVP-51
27	128	74.0	36	21	AA16167	HIV isolate WVP-51
28	128	74.0	36	21	AA16167	HIV isolate WVP-51
29	128	74.0	36	21	AA16167	HIV isolate WVP-51
30	128	74.0	167	22	AA16167	HIV isolate WVP-51
31	128	74.0	204	19	AA16167	HIV isolate WVP-51
32	128	74.0	351	15	AA16167	HIV isolate WVP-51
33	128	74.0	351	15	AA16167	HIV isolate WVP-51
34	127	73.4	32	22	AA16167	HIV isolate WVP-51
35	127	73.4	33	22	AA16167	HIV isolate WVP-51
36	127	73.4	36	20	AA16167	HIV isolate WVP-51
37	127	73.4	36	21	AA16167	HIV isolate WVP-51
38	127	73.4	36	21	AA16167	HIV isolate WVP-51
39	127	73.4	36	21	AA16167	HIV isolate WVP-51
40	127	73.4	36	21	AA16167	HIV isolate WVP-51
41	127	73.4	36	21	AA16167	HIV isolate WVP-51
42	127	73.4	40	17	AA16167	HIV isolate WVP-51
43	127	73.4	111	20	AA16167	HIV isolate WVP-51
44	126	72.8	33	21	AA12230	Partial sequence o
45	126	72.8	40	17	AA12230	Partial sequence o

#### ALIGNMENTS

RESULT 1	AA12264	standard; peptide: 30 AA.
ID	AA12264	standard; peptide: 30 AA.
AC	AA12264	
DT	10-NOV-2000	(first entry)
XX		
DE	HIV-1 gp41 immunodominant region consensus sequence peptide 147-5.	
XX		
KW	HIV-1; AIDS: human immunodeficiency virus type 1; antibody detection; acquired immunodeficiency syndrome; group O HIV; gp41.	
XX		
OS	Human immunodeficiency virus type 1.	
XX		
PN	EP1013766-A2.	
XX		
PD	28-JUN-2000.	
XX		
PF	29-NOV-1999.	99EP-0309491.
XX		
PR	30-NOV-1998.	98US-0110292.
PR	08-FEB-1999.	99US-0119138.
PR	04-NOV-1999.	99US-0433428.
XX		
PA	(ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.	
XX		
PI	De Leys R, Zheng J;	
XX		
DR	WPI: 2000-402205/35.	
XX		
PT	New antigenic peptides and peptide functional derivatives, useful for detection of antibodies produced in response to human immunodeficiency virus group O antibodies -	

PA (SNFI) PASTEUR SANO. PTCS SA.  
 XX Chenebaux DMB, Delagneau H, Gadelles SJX, Rleunier FX;  
 XX WPI: 1998-583190/49.  
 DR  
 XX New synthetic peptide(s) - useful for, e.g. detecting infection by  
 PT human immune deficiency virus of group O  
 XX  
 PS Claim 6; Page 44; 55pp; French.  
 XX  
 CC AAM80459-74 represent synthetic peptides (either linear or cyclised by  
 CC Cys-Cys disulphide bonds). The peptides represent variable sequences  
 CC connected around short highly conserved sequences present in isolates  
 CC of group O human immune deficiency virus (HIV). The peptides are  
 CC useful as immunological reagents for detecting infection by group O  
 CC human immune deficiency virus (HIV).  
 CC  
 SQ Sequence 32 AA;

Query Match 75.1%; Score 130; DB 19; Length 32;  
 Best Local Similarity 71.4%; Pred. No. 1,1e-11;  
 Matches 20; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSMCGKGRITCYTSARWH 30  
 |||:||||:||||:||||:||||:|  
 DB 3 etlmgqqlnswgcgrtlvcysvrvwn 30

# RESULT 14

AAB12229 AAB12229 standard; peptide: 33 AA.

AC AAB12229;

DT 10-NOV-2000 (first entry)

DE Partial sequence of HIV-1 strain HLD28 gp41 immunodominant region.

XX HIV-1; AIDS; human immunodeficiency virus type 1; immunodominant region;  
 KM acquired immunodeficiency syndrome; group O HIV; gp41; HLD28.

OS Human immunodeficiency virus type 1.  
 EP1013766-A2.

XX EP1013766-A2.

PD 28-JUN-2000.

PF 29-NOV-1999; 99EP-0309491.

PR 30-NOV-1998; 98US-0110292.

PR 08-FEB-1999; 99US-0119136.

PR 04-NOV-1999; 99US-0433428.

PA (ORTH-) ORTHO-CLINICAL DIAGNOSTICS INC.

PI De Leye R, Zheng J;

DR WPI: 2000-402205/35.

PT New antigenic peptides and peptide functional derivatives, useful for  
 PT detection of antibodies produced in response to human immunodeficiency  
 PT virus group O antibodies -

XX Example 1; Fig 1; 52pp; English.

CC The present sequence is a partial gp41 protein of Human Immunodeficiency  
 CC Virus Type 1 (HIV-1) strain HLD28. HIV is the principle aetiological  
 CC agent for acquired immunodeficiency syndrome (AIDS). gp41 is a HIV  
 CC envelope protein, and may be used as an antigen for the detection of  
 CC antibodies produced in response to HIV infection. HLD28 is a member of  
 CC HIV group O (outlier). The present sequence is the immunodominant region

CC of gp41. This sequence was used in a sequence homology alignment,  
 CC which in turn was used to derive a consensus sequence peptide: peptide  
 CC 147 (AAB12254).  
 CC  
 SQ Sequence 33 AA;

Query Match 75.1%; Score 130; DB 21; Length 33;  
 Best Local Similarity 75.0%; Pred. No. 1,1e-11;  
 Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSMCGKGRITCYTSARWH 30  
 |||:||||:||||:||||:||||:|  
 DB 6 etlmgqqlnswgcgrtlvcysvrvwn 33

# RESULT 15

AAV67606 AAV67606 standard; peptide: 36 AA.

AC AAV67606;

DT 23-MAR-2000 (first entry)

DE Peptide #6 for detecting HIV-1 group O infection.

KM Human immunodeficiency virus-1; HIV-1; gp41 envelope protein; detection;  
 KM Increased structural stability; diagnostic antigen.

OS Synthetic.

PN WO962945-A2.

PD 09-DEC-1999.

PF 04-JUN-1999; 99WO-US12446.

PR 05-JUN-1998; 98US-0088229.

PR 15-SEP-1998; 98US-0098705.

PR 28-JAN-1999; 99WO-US01726.

PA (PEPT-) PEPTIDE SOLUTIONS INC.

PI Chowdhury MA, Bernstein D, Molsenbocker MA;

DR WPI: 2000-086953/07.

PT Improving properties of peptides for use as diagnostic antigens or for  
 PT preventing or treating infections -

XX Claim 18; Page 61; 83pp; English.

CC This peptide is derived from the human immunodeficiency virus (HIV)-1  
 CC gp41 envelope protein, and is useful for detecting HIV-1 group O  
 CC infection. The invention relates to peptides derived from HIV-1 which  
 CC have been modified for use as diagnostic antigens in the treatment or  
 CC prevention of infection. The structural stability of the peptides can be  
 CC increased in four different ways: through the replacement of a  
 CC hydrophobic amino acid with a less hydrophobic amino acid; through an  
 CC increase in the amount of secondary structure (i.e. alpha helix) in the  
 CC peptide; through the removal of a positive charge from the peptide, or  
 CC through the constraint of the epitopic sequence via the formation of a  
 CC covalent crosslink. Modified peptides of the invention are used to detect  
 CC infectious agents specifically HIV-1. Other detectable agents include  
 CC Group O viruses; human T-cell lymphotropic virus-I or -II; hepatitis C  
 CC and the causative agent of syphilis. The peptides can be used for  
 CC prevention or treatment of infections (e.g. as vaccines, or where  
 CC expressed from a transgene). More generally almost any peptide can be  
 CC similarly modified, e.g. cytokines and interferons; major  
 CC histocompatibility complex antigens; hormones; growth factors; tumour  
 CC markers or suppressors, or antigens from many other pathogens.

Sequence 36 AA;

Query Match 75.1%; Score 130; DB 21; Length 36;  
 Best Local Similarity 75.0%; Pred. No. 1.3e-11;  
 Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 ETLMDODORLNSGCKGRRICTYSARWH 30  
 |||:||||| |||||:||||| :||  
 Db 8 etllqngqrlnlwgcckgrllcyltslkn 35

Search completed: June 20, 2002, 15:31:31  
 Job time: 314 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 20, 2002, 15:28:12 ; Search time 12.97 Seconds

(without alignments)  
56.497 Million cell updates/sec

Title: US-09-605-573A-69

Perfect score: 173  
Sequence: 1 GRETLMODQRLNSGCKGRITCYTSARWH 30Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents, AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	173	100.0	30	4	US-09-433-428D-69
2	166	96.0	30	4	US-09-433-428D-63
3	157	90.8	149	4	US-09-433-428D-67
4	157	90.8	220	4	US-09-433-428D-66
5	157	90.8	368	4	US-09-433-428D-58
6	157	90.8	439	4	US-09-433-428D-57
7	155	89.6	28	4	US-09-433-428D-62
8	155	89.6	35	4	US-09-433-428D-64
9	134	77.5	23	4	US-09-433-428D-61
10	133	76.9	23	4	US-09-433-428D-59
11	132	76.3	23	4	US-09-433-428D-60
12	131	75.7	33	4	US-09-433-428D-30
13	131	75.7	40	3	US-08-894-699-36
14	131	75.7	40	4	US-09-444-410-36
15	130	75.1	33	4	US-09-433-428D-23
16	129	74.6	33	4	US-09-433-428D-25
17	129	74.6	40	3	US-08-894-699-39
18	129	74.6	40	4	US-09-444-410-39
19	128	74.0	33	4	US-09-433-428D-2
20	128	74.0	33	4	US-09-433-428D-9
21	128	74.0	33	4	US-09-433-428D-10
22	128	74.0	33	4	US-09-433-428D-12
23	128	74.0	33	4	US-09-433-428D-13
24	128	74.0	35	1	US-08-470-202-62
25	128	74.0	35	1	US-08-471-770-62
26	128	74.0	35	2	US-08-468-059-62
27	128	74.0	35	4	US-09-109-916-62

28	128	74.0	146	2	US-08-394-021-10	Sequence 10, Appl
29	128	74.0	146	4	US-09-131-551-10	Sequence 10, Appl
30	128	74.0	204	4	US-08-965-056-105	Sequence 105, Appl
31	128	74.0	351	1	US-08-470-202-46	Sequence 46, Appl
32	128	74.0	351	1	US-08-471-770-46	Sequence 46, Appl
33	128	74.0	351	2	US-08-468-059-46	Sequence 46, Appl
34	128	74.0	351	4	US-09-109-916-46	Sequence 46, Appl
35	127	73.4	33	4	US-09-433-428D-6	Sequence 6, Appl
36	127	73.4	40	3	US-08-894-699-68	Sequence 68, Appl
37	127	73.4	40	4	US-09-444-410-68	Sequence 68, Appl
38	126	72.8	33	4	US-09-433-428D-24	Sequence 24, Appl
39	126	72.8	40	3	US-08-894-699-38	Sequence 38, Appl
40	126	72.8	40	4	US-09-444-410-38	Sequence 38, Appl
41	125	72.3	33	4	US-09-433-428D-20	Sequence 20, Appl
42	125	72.3	40	3	US-08-894-699-41	Sequence 41, Appl
43	125	72.3	40	4	US-09-444-410-41	Sequence 41, Appl
44	125	72.3	356	1	US-08-602-713-12	Sequence 12, Appl
45	125	72.3	356	4	US-08-989-493-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
US-09-433-428D-69

; Sequence 69, Application US/09433428D

; Patent No. 6149910

Query Match 100.0%; Score 173; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.2e-19;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRETLMODQRLNSGCKGRITCYTSARWH 30  
DB 1 GRETLMODQRLNSGCKGRITCYTSARWH 30

RESULT 2  
US-09-433-428D-63  
; Sequence 63, Application US/09433428D  
; Patent No. 6149910  
; GENERAL INFORMATION:  
; APPLICANT: De Leys, Robert J.  
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
; FILE REFERENCE: CDS-207  
; CURRENT APPLICATION NUMBER: US/09/433,428D  
; CURRENT FILING DATE: 1999-11-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 63  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:

NAME/KEY: Xaa is any amino acid  
LOCATION: 8  
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
OTHER INFORMATION: Sequence  
US-09-433-428D-63

Query Match 96.0%; Score 166; DB 4; Length 30;  
Best Local Similarity 96.7%; Pred. No. 2.4e-18;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRTLMQDOORLNSWCGCKGRITCYTSARWH 30  
|||||:|||||:|||||:|||||:|||||  
Db 1 GRTLMQXOORLNSWCGCKGRITCYTSARWH 30

RESULT 3  
US-09-433-428D-67  
Sequence 67, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433.428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 67  
LENGTH: 149  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
US-09-433-428D-67

Query Match 90.8%; Score 157; DB 4; Length 149;  
Best Local Similarity 96.4%; Pred. No. 3.1e-16;  
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 ETLMOQOORLNSWCGCKGRITCYTSARWH 30  
|||||:|||||:|||||:|||||:|||||  
Db 44 ETLMOQOORLNSWCGCKGRITCYTSARWH 71

RESULT 4  
US-09-433-428D-66  
Sequence 66, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433.428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 220  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
US-09-433-428D-66

Query Match 90.8%; Score 157; DB 4; Length 220;  
Best Local Similarity 96.4%; Pred. No. 4.8e-16;

Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 ETLMOQOORLNSWCGCKGRITCYTSARWH 30  
|||||:|||||:|||||:|||||:|||||  
Db 44 ETLMOQOORLNSWCGCKGRITCYTSARWH 71

RESULT 5  
US-09-433-428D-58  
Sequence 58, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433.428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 58  
LENGTH: 368  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
US-09-433-428D-58

Query Match 90.8%; Score 157; DB 4; Length 368;  
Best Local Similarity 96.4%; Pred. No. 8.5e-16;  
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 ETLMOQOORLNSWCGCKGRITCYTSARWH 30  
|||||:|||||:|||||:|||||:|||||  
Db 240 ETLMOQOORLNSWCGCKGRITCYTSARWH 267

RESULT 6  
US-09-433-428D-57  
Sequence 57, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433.428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 57  
LENGTH: 439  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus  
US-09-433-428D-57

Query Match 90.8%; Score 157; DB 4; Length 439;  
Best Local Similarity 96.4%; Pred. No. 1e-15;  
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 ETLMOQOORLNSWCGCKGRITCYTSARWH 30  
|||||:|||||:|||||:|||||:|||||  
Db 240 ETLMOQOORLNSWCGCKGRITCYTSARWH 267

RESULT 7  
US-09-433-428D-62

```
Sequence 62, Application US/09433428D
Patent No. 6149910
GENERAL INFORMATION:
APPLICANT: De Leys, Robert J.
APPLICANT: Zheng, Jian
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
FILE REFERENCE: CDS-207
CURRENT APPLICATION NUMBER: US/09/433,428D
CURRENT FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 62
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Xaa is any amino acid
LOCATION: 6
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
OTHER INFORMATION: Sequence
US-09-433-428D-62
```

```
Query Match      89.6%; Score 155; DB 4; Length 28;
Best Local Similarity 96.4%; Pred. No. 1e-16;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 3 ETLMDODRLNSMCGKGRITCYTSARWH 30
    ||||| ||||| ||||| ||||| |||||
Db 1 ETLMDODRLNSMCGKGRITCYTSARWH 28
```

```
RESULT 8
US-09-433-428D-64
Sequence 64, Application US/09433428D
Patent No. 6149910
GENERAL INFORMATION:
APPLICANT: De Leys, Robert J.
APPLICANT: Zheng, Jian
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
FILE REFERENCE: CDS-207
CURRENT APPLICATION NUMBER: US/09/433,428D
CURRENT FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 64
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Xaa is any amino acid
LOCATION: 13
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
OTHER INFORMATION: Sequence
US-09-433-428D-64
```

```
Query Match      89.6%; Score 155; DB 4; Length 35;
Best Local Similarity 96.4%; Pred. No. 1.3e-16;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 3 ETLMDODRLNSMCGKGRITCYTSARWH 30
    ||||| ||||| ||||| ||||| |||||
Db 8 ETLMDODRLNSMCGKGRITCYTSARWH 35
```

```
RESULT 9
US-09-433-428D-61
Sequence 61, Application US/09433428D
Patent No. 6149910
GENERAL INFORMATION:
APPLICANT: De Leys, Robert J.
APPLICANT: Zheng, Jian
```

```
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
FILE REFERENCE: CDS-207
CURRENT APPLICATION NUMBER: US/09/433,428D
CURRENT FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 61
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
OTHER INFORMATION: Sequence
US-09-433-428D-61
```

```
Query Match      77.5%; Score 134; DB 4; Length 23;
Best Local Similarity 95.7%; Pred. No. 1.1e-13;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 8 DQORLNSMCGKGRITCYTSARWH 30
    :|||: ||||| ||||| ||||| |||||
Db 1 EQORLNSMCGKGRITCYTSARWH 23
```

```
RESULT 10
US-09-433-428D-59
Sequence 59, Application US/09433428D
Patent No. 6149910
GENERAL INFORMATION:
APPLICANT: De Leys, Robert J.
APPLICANT: Zheng, Jian
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
FILE REFERENCE: CDS-207
CURRENT APPLICATION NUMBER: US/09/433,428D
CURRENT FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 59
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
OTHER INFORMATION: Sequence
US-09-433-428D-59
```

```
Query Match      76.9%; Score 133; DB 4; Length 23;
Best Local Similarity 95.7%; Pred. No. 1.6e-13;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 8 DQORLNSMCGKGRITCYTSARWH 30
    :|||: ||||| ||||| ||||| |||||
Db 1 NOORLNSMCGKGRITCYTSARWH 23
```

```
RESULT 11
US-09-433-428D-60
Sequence 60, Application US/09433428D
Patent No. 6149910
GENERAL INFORMATION:
APPLICANT: De Leys, Robert J.
APPLICANT: Zheng, Jian
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
FILE REFERENCE: CDS-207
CURRENT APPLICATION NUMBER: US/09/433,428D
CURRENT FILING DATE: 1999-11-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 60
LENGTH: 23
TYPE: PRT
```

ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: Xaa is any amino acid  
LOCATION: 1  
OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus  
US-09-433-428D-60

Query Match  
Best Local Similarity 76.3%; Score 132; DB 4; Length 23;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 QOURLNSWGCKGRICVTSARWH 30  
|||||  
DB 2 QOURLNSWGCKGRICVTSARWH 23

RESULT 12  
US-09-433-428D-30  
Sequence 30, Application US/09433428D  
Patent No. 6149910  
GENERAL INFORMATION:  
APPLICANT: De Leys, Robert J.  
APPLICANT: Zheng, Jian  
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O  
FILE REFERENCE: CDS-207  
CURRENT APPLICATION NUMBER: US/09/433,428D  
CURRENT FILING DATE: 1999-11-04  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-433-428D-30

Query Match  
Best Local Similarity 75.7%; Score 131; DB 4; Length 33;  
Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 ETLMQOQRLNSWGCKGRICVTSARWH 30  
|||||  
DB 6 ETLMQOQRLNSWGCKGRICVTSARWH 33

RESULT 13  
US-08-894-699-36  
Sequence 36, Application US/08894699  
Patent No. 6030769  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSEY-ATAKA, IBITISSAM  
APPLICANT: LY, THOI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,699  
FILING DATE: 01-DEC-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR96/00294  
FILING DATE: 26-FEB-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236  
FILING DATE: 27-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,614  
REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-699-36

Query Match  
Best Local Similarity 75.7%; Score 131; DB 3; Length 40;  
Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 ETLMQOQRLNSWGCKGRICVTSARWH 30  
|||||  
DB 11 ETLMQOQRLNSWGCKGRICVTSARWH 38

RESULT 14  
US-09-444-410-36  
Sequence 36, Application US/09444410  
Patent No. 6270975  
GENERAL INFORMATION:  
APPLICANT: SIMON, FRANCOIS  
APPLICANT: SARAGOSTI, SENTOB  
APPLICANT: LOUSSEY-ATAKA, IBITISSAM  
APPLICANT: LY, THOI-DUONG  
APPLICANT: CHAIX-BAUDIER, MARIE-LAURE  
TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,410  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/894,699  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: FR 95/02236

? FILING DATE: 27-FEB-1995  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: OBLON, NORMAN F.  
 ? REGISTRATION NUMBER: 24,614  
 ? REFERENCE/DOCKET NUMBER: 0354-0020-0PCT  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: 703-413-3000  
 ? TELEFAX: 703-413-2220  
 ? INFORMATION FOR SEQ ID NO: 36:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 40 amino acids  
 ? TYPE: amino acid  
 ? STRANDEDNESS: single  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: peptide  
 ? US-09-444-410-36

Query Match	75.7%	Score 131	DB 4	length 40
Best Local Similarity	75.0%	Pred. No.	5.7e-13	
Matches	21	Conservative	5	Mismatches 2; Indels 0; Gaps 0;

```
QY      3 ETLMDQDORLNSWGCKGRITCYTSARWH 30
          |||::|| | |||||::||| : |:
Db     11 ETLIQNLNSWGCGRIVCYTSVKWN 38
```

```

RESULT 15 428D-23
US-09-433-428D-23
: Sequence 23, Application US/09433428D
: Patent No. 6149810
: GENERAL INFORMATION:
: APPLICANT: De Leys, Robert J.
: APPLICANT: Zheng, Jian
: TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
: FILE REFERENCE: CDS-207
: CURRENT APPLICATION NUMBER: US/09/433,428D
: CURRENT FILING DATE: 1999-11-04
: NUMBER OF SEQ ID NOS: 70
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 23
: LENGTH: 33
: TYPE: PRT
: ORGANISM: Human Immunodeficiency virus type 1
US-09-433-428D-23

```

Query Match	75.1%	Score 130;	DB 4;	Length 33;
Best Local Similarity	75.0%	Pred. No. 6.5e-13;		
Matches 21;	Conservative 5;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      3  ETLMQDQRLNSWGCKGRITCYTSARWH  30
          |||||:|||||:|||||:|:
Db      6  ETLMQDQRLDLWGCKGRITCYTSVKNW  33
          .
```

Search completed: June 20, 2002, 15:31:50  
Job time: 218 sec

THIS PAGE BLANK (18470)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## \* OM protein - protein search, using sw model

Run on: June 20, 2002, 15:30:32 ; Search time 105.35 Seconds

(without alignments)  
100.232 Million cell updates/sec

Title: US-09-605-573a-69

Perfect score: 173  
Sequence: 1 GRETLMODQRLNSWCGKGRICYSARWH 30Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	173	100.0	30	18	US-09-433-428-69
2	173	100.0	30	20	US-09-605-573a-69
3	166	96.0	30	18	US-09-433-428-63
4	166	96.0	30	20	US-09-605-573a-63
5	157	90.8	149	18	US-09-433-428-67
6	157	90.8	149	20	US-09-605-573a-67
7	157	90.8	220	18	US-09-433-428-66

	8	157	90.8	220	20	US-09-605-573a-66	Sequence 66, Appl
	9	157	90.8	368	18	US-09-433-428-58	Sequence 58, Appl
	10	157	90.8	368	20	US-09-605-573a-58	Sequence 58, Appl
	11	157	90.8	439	18	US-09-433-428-57	Sequence 57, Appl
	12	157	90.8	439	20	US-09-605-573a-57	Sequence 57, Appl
	13	155	89.6	28	18	US-09-433-428-62	Sequence 62, Appl
	14	155	89.6	28	20	US-09-605-573a-62	Sequence 62, Appl
	15	155	89.6	35	18	US-09-433-428-64	Sequence 64, Appl
	16	155	89.6	35	20	US-09-605-573a-64	Sequence 64, Appl
	17	134	77.5	23	18	US-09-433-428-61	Sequence 61, Appl
	18	134	77.5	23	20	US-09-605-573a-61	Sequence 61, Appl
	19	133	76.9	23	18	US-09-433-428-59	Sequence 59, Appl
	20	133	76.9	23	20	US-09-605-573a-59	Sequence 59, Appl
	21	132	76.3	23	18	US-09-433-428-60	Sequence 60, Appl
	22	132	76.3	23	20	US-09-605-573a-60	Sequence 60, Appl
	23	131	75.7	33	18	US-09-433-428-30	Sequence 30, Appl
	24	131	75.7	33	20	US-09-605-573a-30	Sequence 30, Appl
	25	130	75.1	32	15	US-09-147-362-11	Sequence 11, Appl
	26	130	75.1	32	18	US-09-147-362a-11	Sequence 11, Appl
	27	130	75.1	33	18	US-09-433-428-23	Sequence 23, Appl
	28	130	75.1	33	20	US-09-605-573a-23	Sequence 23, Appl
	29	130	75.1	113	18	US-09-462-917a-18	Sequence 18, Appl
	30	129	74.6	33	18	US-09-433-428-25	Sequence 25, Appl
	31	129	74.6	33	20	US-09-605-573a-25	Sequence 25, Appl
	32	128	74.0	33	18	US-09-433-428-2	Sequence 2, Appl
	33	128	74.0	33	20	US-09-605-573a-2	Sequence 2, Appl
	34	128	74.0	33	18	US-09-433-428-10	Sequence 10, Appl
	35	128	74.0	33	20	US-09-605-573a-10	Sequence 10, Appl
	36	128	74.0	33	18	US-09-433-428-12	Sequence 12, Appl
	37	128	74.0	33	20	US-09-605-573a-12	Sequence 12, Appl
	38	128	74.0	33	18	US-09-433-428-13	Sequence 13, Appl
	39	128	74.0	33	20	US-09-605-573a-9	Sequence 9, Appl
	40	128	74.0	33	20	US-09-605-573a-10	Sequence 10, Appl
	41	128	74.0	33	20	US-09-605-573a-12	Sequence 12, Appl
	42	128	74.0	33	15	US-09-147-362-21	Sequence 21, Appl
	43	128	74.0	35	15	US-09-147-362a-21	Sequence 21, Appl
	44	128	74.0	35	22	US-09-886-149-62	Sequence 62, Appl
	45	128	74.0	35	22	US-09-886-150-62	Sequence 62, Appl

## ALIGNMENTS

RESULT 1

US-09-433-428-69

Sequence 69, Application US/09433428

GENERAL INFORMATION:

APPLICANT: Zheng, Jian

FILE REFERENCE: CDS-207

CURRENT APPLICATION NUMBER: US/09/433,428

CURRENT FILING DATE: 1999-11-04

NUMBER OF SEQ ID NOS: 69

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 69

LENGTH: 30

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus

OTHER INFORMATION: Sequence

US-09-433-428-69

Query Match 100.0%; Score 173; DB 18; Length 30;  
Best Local Similarity 100.0%; Pred. No. 7e-17;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GRETLMODQRLNSWCGKGRICYSARWH 30

1 GRETLMODQRLNSWCGKGRICYSARWH 30

```
RESULT 2
US-09-605-573a-69
; Sequence 69, Application US/09605573A
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-222
; CURRENT APPLICATION NUMBER: US/09/605,573A
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 69
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
; OTHER INFORMATION: Sequence
US-09-605-573a-69

Query Match          100.0%; Score 173; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRETLMQDOORLNSWCGCKGRILICYTSARWH 30
Db 1 GRETLMQDOORLNSWCGCKGRILICYTSARWH 30

RESULT 3
US-09-433-428-63
; Sequence 63, Application US/09433428
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 63
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
; OTHER INFORMATION: Sequence
US-09-433-428-63

Query Match          96.0%; Score 166; DB 18; Length 30;
Best Local Similarity 96.7%; Pred. No. 6.9e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRETLMQDOORLNSWCGCKGRILICYTSARWH 30
Db 1 GRETLMQDOORLNSWCGCKGRILICYTSARWH 30

RESULT 4
US-09-605-573a-63
; Sequence 63, Application US/09605573A
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-222
; CURRENT APPLICATION NUMBER: US/09/605,573A
; CURRENT FILING DATE: 2000-06-28
```

```
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 63
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 8
; OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
; OTHER INFORMATION: Sequence
US-09-605-573a-63

Query Match          96.0%; Score 166; DB 20; Length 30;
Best Local Similarity 96.7%; Pred. No. 6.9e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRETLMQDOORLNSWCGCKGRILICYTSARWH 30
Db 1 GRETLMQDOORLNSWCGCKGRILICYTSARWH 30

RESULT 5
US-09-433-428-67
; Sequence 67, Application US/09433428
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 67
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
; OTHER INFORMATION: Sequence
US-09-433-428-67

Query Match          90.8%; Score 157; DB 18; Length 149;
Best Local Similarity 96.4%; Pred. No. 6.2e-14;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 ETLMOQDOORLNSWCGCKGRILICYTSARWH 30
Db 44 ETLMOQDOORLNSWCGCKGRILICYTSARWH 71

RESULT 6
US-09-605-573a-67
; Sequence 67, Application US/09605573A
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-222
; CURRENT APPLICATION NUMBER: US/09/605,573A
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 67
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
; OTHER INFORMATION: Sequence
```

US-09-605-573A-67

Query Match	90.88;	Score 157;	DB 20;	Length 149;
Best Local Similarity	96.48;	Pred. No. 6.2e-14;		
Matches 27; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      3  ETLMDQDQRLNSWGCKGRILCYTSARWH 30
          |||||:|||||
Db      44  ETLMDQDQRLNSWGCKGRILCYTSARWH 71
```

RESULT 7  
US-09-433-428-66

```

: GENERAL INFORMATION:
: APPLICANT: De Leys, Robert J.
: APPLICANT: zheng, jian
: TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
: FILE REFERENCE: CDS-207
: CURRENT APPLICATION NUMBER: US/09/433.428
: CURRENT FILING DATE: 1999-11-04
: NUMBER OF SEQ ID NOS: 69
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 66
: LENGTH: 220
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
: OTHER INFORMATION: Sequence
: US-09-433-428-66

```

Query Match	90.8%	Score 157	DB 18	Length 220
Best Local Similarity	96.4%	Pred.No. 9e-14		
Matches 27	Conservative 1	Mismatches 0	Indels 0	Gaps 0

```

Oy      3  ETLMDODQRLNSWGCGRICYTSARWH 30
          ||||:|||||||
Db      44  ETLMDNQRLNSWGCGRICYTSARWH 71

```

RESULT 8  
US-09-605-573A-66

```

: GENERAL INFORMATION:
: APPLICANT: De Leys, Robert J.
: APPLICANT: Zheng, Jian
: TITLE OR INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
: FILE REFERENCE: CDS-222
: CURRENT APPLICATION NUMBER: US/09/605,573A
: CURRENT FILING DATE: 2000-06-28
: NUMBER OF SEQ ID NOS: 70
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 66
: LENGTH: 220
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: HIV Consensus
: OTHER INFORMATION: Sequence
: US-09-605-573A-66

```

Query Match	90.8%	Score 157	DB 20	Length 220
Best Local Similarity	96.4%	Pred. No. 9e-14		
Matches 27	Conservative 1	Mismatches 0	Indels 0	Gaps 0

```

Oy      3  ETLMDOOORLNSWGCKGRIICYTSARWH 30
         |||:|||||
Db      4 4  ETLMONQORLNSWGCKGRIICYTSARWH 71

```

RESULT 9  
US-09-433-428-58

```

? GENERAL INFORMATION:
? APPLICANT: De leys, Robert J.
? APPLICANT: Zheng, Jian
? TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
? FILE REFERENCE: CDS-207
? CURRENT APPLICATION NUMBER: US/09/433,428
? CURRENT FILING DATE: 1999-11-04
? NUMBER OF SEQ ID NOS: 69
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 58
? LENGTH: 368
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
? OTHER INFORMATION: Sequence
US-09-433-428-58

```

Query Match	90.8%;	Score 157;	DB 18;	Length 368;
Best Local Similarity	96.4%;	Pred. NO. 1.5e-13;		
Matches 27; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0

QY	3	ETLMQDQRLNSWGCKGRIICYTSARWH	30
		:	
Db	240	ETLMQDQRLNSWGCKGRIICYTSARWH	267

RESULT 10  
US-09-605-573A-58  
; sequence 58, Application US/09605573A

```

GENERAL INFORMATION:
APPLICANT: De Leys, Robert J.
APPLICANT: zheng, jian
TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
FILE REFERENCE: CDS-222
CURRENT APPLICATION NUMBER: US/09/605,573A
CURRENT FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 58
LENGTH: 368
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
OTHER INFORMATION: Sequence
US-09-605-573A-58

```

Query Match	90.88;	Score 157;	DB 20;	Length 368;
Best Local Similarity	96.48;	Pred. No. 1.5e-13;		
Matches 27; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0

Oy	3	ETLMDDQRLNSWGCKGRICYSARWH	30
		:	
Db	240	ETLMQNDQRLNSWGCKGRICYSARWH	267

RESULT 11  
US-09-433-428-57

```

: Sequence 57, Application US/09/433,428
: GENERAL INFORMATION:
: APPLICANT: De leys, Robert J.
: APPLICANT: Zheng, Jian
: TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP C
: FILE REFERENCE: CDS-207
: CURRENT APPLICATION NUMBER: US/09/433,428
: CURRENT FILING DATE: 1999-11-04

```

```

; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
US-09-433-428-57
```

```
Query Match          90.8%; Score 157; DB 18; Length 439;
Best Local Similarity 96.4%; Pred. No. 1.8e-13;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 ETLMOQOORLNSWGCKGRITCYTSARWH 30
|||||:|||||:|||||:|||||:|||||
Db 240 ETLMOQOORLNSWGCKGRITCYTSARWH 267
```

```
RESULT 12
US-09-605-573A-57
; Sequence 57, Application US/09605573A
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-222
; CURRENT APPLICATION NUMBER: US/09/605,573A
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
US-09-605-573A-57
```

```
Query Match          90.8%; Score 157; DB 20; Length 439;
Best Local Similarity 96.4%; Pred. No. 1.8e-13;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3 ETLMOQOORLNSWGCKGRITCYTSARWH 30
|||||:|||||:|||||:|||||:|||||
Db 240 ETLMOQOORLNSWGCKGRITCYTSARWH 267
```

```
RESULT 13
US-09-433-428-62
; Sequence 62, Application US/09433428
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
US-09-433-428-62
```

```
Query Match          89.6%; Score 155; DB 18; Length 28;
Best Local Similarity 96.4%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 3 ETLMOQOORLNSWGCKGRITCYTSARWH 30
|||||:|||||:|||||:|||||:|||||
Db 1 ETLMOQOORLNSWGCKGRITCYTSARWH 28
```

```
RESULT 14
US-09-605-573A-62
; Sequence 62, Application US/09605573A
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-222
; CURRENT APPLICATION NUMBER: US/09/605,573A
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Xaa is any amino acid
; LOCATION: 6
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
US-09-605-573A-62
```

```
Query Match          89.6%; Score 155; DB 20; Length 28;
Best Local Similarity 96.4%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 3 ETLMOQOORLNSWGCKGRITCYTSARWH 30
|||||:|||||:|||||:|||||:|||||
Db 1 ETLMOQOORLNSWGCKGRITCYTSARWH 28
```

```
RESULT 15
US-09-433-428-64
; Sequence 64, Application US/09433428
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert J.
; APPLICANT: Zheng, Jian
; TITLE OF INVENTION: PEPTIDES FOR THE DETECTION OF HIV-1 GROUP O
; FILE REFERENCE: CDS-207
; CURRENT APPLICATION NUMBER: US/09/433,428
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:HIV Consensus
US-09-433-428-64
```

```
Query Match          89.6%; Score 155; DB 18; Length 35;
Best Local Similarity 96.4%; Pred. No. 3e-14;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 3 ETLMOQOORLNSWGCKGRITCYTSARWH 30
|||||:|||||:|||||:|||||:|||||
Db 8 ETLMOQOORLNSWGCKGRITCYTSARWH 35
```

Fri Jun 21 10:39:05 2002

us-09-605-573a-69.rapm

Page 5

Search completed: June 20, 2002, 15:34:04  
Job time: 212 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 15:31:32 ; Search time 14.82 Seconds  
(without alignments)  
177.533 Million cell updates/sec

Title: US-09-605-573A-69

Perfect score: 173

Sequence: 1 GRETLMDDQRLNSWCKGRICYSARWH 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 265615 seqs, 87701152 residues

Total number of hits satisfying chosen parameters: 265615

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/prodata/1/paa/US06\_NEW\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/paa/US07\_NEW\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/paa/US08\_NEW\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/paa/US60\_NEW\_COMB.pep.\*  
7: /cgn2\_6/prodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	49.1	657	5 US-09-914-205-2	Sequence 2, Appl
2	85	49.1	849	5 US-09-899-575-23	Sequence 23, Appl
3	85	49.1	854	5 US-09-913-159A-2	Sequence 2, Appl
4	85	49.1	855	5 US-09-899-575-24	Sequence 24, Appl
5	85	49.1	867	6 US-10-130-157-363	Sequence 363, App
6	72	41.6	15	4 US-08-776-188B-8	Sequence 8, Appl
7	72	41.6	19	4 US-08-776-188B-19	Sequence 19, Appl
8	72	41.6	19	4 US-08-776-188B-36	Sequence 36, Appl
9	72	41.6	22	5 US-09-576-824A-20	Sequence 19, Appl
10	72	41.6	27	4 US-08-776-188B-10	Sequence 10, Appl
11	72	41.6	31	4 US-08-776-188B-54	Sequence 54, Appl
12	70	40.5	26	5 US-09-359-575-14	Sequence 14, Appl
13	65	37.6	23	4 US-08-776-188B-5	Sequence 5, Appl
14	65	37.6	24	4 US-09-576-824A-6	Sequence 6, Appl
15	55	32.1	20	4 US-08-776-188B-30	Sequence 30, Appl
16	55	31.8	20	4 US-08-776-188B-48	Sequence 48, Appl
17	55	31.5	344	6 US-10-106-698-5894	Sequence 5894, Ap
18	54.5	31.5	16	4 US-08-776-188B-39	Sequence 39, Appl
19	54	31.2	18	4 US-08-776-188B-53	Sequence 53, Appl
20	53	30.6	12	5 US-09-576-824A-5	Sequence 5, Appl
21	53	30.6	20	5 US-09-576-824A-2	Sequence 2, Appl
22	52	30.1	21	4 US-08-776-188B-50	Sequence 50, Appl
23	52	30.1	22	4 US-08-776-188B-33	Sequence 33, Appl
24	52	30.1	22	4 US-08-776-188B-35	Sequence 35, Appl
25	52	30.1	24	4 US-08-776-188B-49	Sequence 49, Appl
26	51.5	29.8	21	5 US-09-576-824A-3	Sequence 3, Appl

27	51	29.5	275	7 US-60-360-039-4600	Sequence 4600, Ap
28	51	29.5	275	7 US-60-360-039-7357	Sequence 7357, Ap
29	49	28.3	12	5 US-09-576-824A-110	Sequence 110, App
30	49	28.3	13	5 US-09-576-824A-177	Sequence 177, App
31	49	28.3	14	5 US-09-576-824A-176	Sequence 176, App
32	49	28.3	14	5 US-09-701-432-9	Sequence 9, Appl
33	49	28.3	15	5 US-09-576-824A-111	Sequence 111, App
34	49	28.3	15	5 US-09-576-824A-111	Sequence 111, App
35	48	27.7	220	6 US-10-138-098-34	Sequence 34, Appl
36	48	27.7	552	7 US-60-360-039-22048	Sequence 22048, A
37	48	27.7	641	6 US-10-138-098-38	Sequence 38, Appl
38	48	27.7	641	6 US-10-138-098-39	Sequence 39, Appl
39	48	27.7	849	7 US-60-360-039-2688	Sequence 2688, Ap
40	48	27.7	1006	6 US-10-138-098-45	Sequence 45, Appl
41	48	27.7	1006	6 US-10-138-098-47	Sequence 46, Appl
42	48	27.7	1006	6 US-10-138-098-47	Sequence 47, Appl
43	48	27.7	1006	6 US-10-138-098-48	Sequence 48, Appl
44	48	27.7	1064	6 US-10-138-098-41	Sequence 41, Appl
45	48	27.7	1064	6 US-10-138-098-42	Sequence 42, Appl

## ALIGNMENTS

RESULT 1  
US-09-914-205-2  
Sequence 2, Application US/09914205  
GENERAL INFORMATION:  
APPLICANT: SIA, Charles D.Y.  
APPLICANT: Cao, Shi-Xian  
APPLICANT: Persson, Roy  
APPLICANT: Rovinsky, Benjamin  
TITLE OF INVENTION: EXPRESSING GP140 FRAGMENT OF PRIMARY HIV-1 ISOLATE  
FILE REFERENCE: 1038-1176 MIS:jb  
CURRENT APPLICATION NUMBER: US/09/914,205  
PRIOR APPLICATION NUMBER: 09/256,194  
PRIOR FILING DATE: 1999-02-24  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 657  
TYPE: PRT  
ORGANISM: Human Immunodeficiency virus type 1  
US-09-914-205-2

Query Match 49.1%; Score 85; DB 5; Length 657;  
Best Local Similarity 46.4%; Pred. No. 0.00033;  
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
OY 3 ETLMDQRLNSWCKGRICYSARWH 30  
DB 565 ERYLDQRLNSWCKGRICYSARWH 592  
RESULT 2  
US-09-899-575-23  
Sequence 23, Application US/09899575  
GENERAL INFORMATION:  
APPLICANT: Zur Megede, Jan  
APPLICANT: Barnett, Susan W.  
APPLICANT: Egnelbrecht, Susan  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
FILE REFERENCE: PP01631.102  
CURRENT APPLICATION NUMBER: US/09/899,575  
CURRENT FILING DATE: 2001-07-05  
PRIOR APPLICATION NUMBER: 09/475,704  
PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 23  
LENGTH: 849  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-899-575-23

Query Match 49.1%; Score 85; DB 5; Length 849;  
Best Local Similarity 46.4%; Pred. No. 0.00043;  
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
Db 577 ERYLKDQDLIGWGCGSKLICITTAVPWN 604

RESULT 3  
US-09-913-159A-2  
Sequence 2, Application US/09913159A  
GENERAL INFORMATION:  
APPLICANT: Strathmann AG & Co.  
TITLE OF INVENTION: Virus-Vaccine  
FILE REFERENCE: P057760  
CURRENT APPLICATION NUMBER: US/09/913,159A  
PRIOR FILING DATE: 2002-03-11  
PRIOR APPLICATION NUMBER: 199 07 485.2  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 854  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
FEATURE:  
OTHER INFORMATION: Envelope Polypeptide  
US-09-913-159A-2

Query Match 49.1%; Score 85; DB 5; Length 854;  
Best Local Similarity 46.4%; Pred. No. 0.00043;  
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
Db 582 ERYLKDQDLIGWGCGSKLICITTAVPWN 609

RESULT 4  
US-09-899-575-24  
Sequence 24, Application US/09899575  
GENERAL INFORMATION:  
APPLICANT: Zur Megecke, Jan  
APPLICANT: Barnett, Susan W.  
APPLICANT: Egnelbrecht, Susan  
APPLICANT: van Rensburg, Estrella Janse  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
FILE REFERENCE: P01631.102  
CURRENT APPLICATION NUMBER: US/09/899,575  
PRIOR FILING DATE: 2001-07-05  
PRIOR APPLICATION NUMBER: 09/475,704  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 855  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-899-575-24

Query Match 49.1%; Score 85; DB 5; Length 855;  
Best Local Similarity 46.4%; Pred. No. 0.00043;

Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
Db 576 ERYLKDQDLIGWGCGSKLICITTAVPWN 603

RESULT 5  
US-10-130-157-363  
Sequence 363, Application US/10130157  
GENERAL INFORMATION:  
APPLICANT: GENPART GmbH  
TITLE OF INVENTION: The Genome of the HIV-1 Inter-subtype (C/B') and Use Thereof  
FILE REFERENCE: WAG-001 PCT  
CURRENT APPLICATION NUMBER: US/10/130,157  
PRIOR FILING DATE: 2002-05-14  
PRIOR APPLICATION NUMBER: DE 19955089.1  
PRIOR FILING DATE: 1999-11-16  
NUMBER OF SEQ ID NOS: 379  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 363  
LENGTH: 867  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-10-130-157-363

Query Match 49.1%; Score 85; DB 6; Length 867;  
Best Local Similarity 46.4%; Pred. No. 0.00044;  
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
Db 588 ERYLKDQDLIGWGCGSKLICITTAVPWN 615

RESULT 6  
US-08-776-188B-8  
Sequence 8, Application US/08776188B  
GENERAL INFORMATION:  
APPLICANT: Weinhuess, Ursula-Henrike  
APPLICANT: Kruse-Mueller, Cornelia  
APPLICANT: Hoss, Eva  
APPLICANT: Faatz, Elke  
APPLICANT: Ofenloch-Hahnel, Beatus  
APPLICANT: Seidel, Christoph  
APPLICANT: Wiedmann, Michael  
TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTI  
FILE REFERENCE: 100564-07003  
CURRENT APPLICATION NUMBER: US/08/776,188B  
PRIOR FILING DATE: 2002-04-15  
PRIOR APPLICATION NUMBER: PCT/EP95/02919  
PRIOR FILING DATE: 1995-07-24  
PRIOR APPLICATION NUMBER: P 44 30 972.4  
PRIOR FILING DATE: 1994-08-31  
PRIOR APPLICATION NUMBER: P 44 26 276.0  
PRIOR FILING DATE: 1994-07-25  
NUMBER OF SEQ ID NOS: 76  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Epitope region of HIV I or  
US-08-776-188B-8

Query Match 41.6%; Score 72; DB 4; Length 15;  
Best Local Similarity 66.7%; Pred. No. 0.00052;  
Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 12 LNSMCKGRIICVYTS 26  
 1: |||||::|||  
 DB 1 LSLMCKGKGLVYTS 15

# RESULT 7

US-08-776-188B-19  
 ; Sequence 19, Application US/08776188B  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Weinhuess, Ursula-Henrike  
 ; APPLICANT: Kruse-Muller, Cornelia  
 ; APPLICANT: Hoss, Eva  
 ; APPLICANT: Faatz, Elke  
 ; APPLICANT: Offenloch-Hahnle, Beatus  
 ; APPLICANT: Seidel, Christoph  
 ; APPLICANT: Wiedmann, Michael  
 ; TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTIGEN  
 ; FILE REFERENCE: 100564-07003  
 ; CURRENT APPLICATION NUMBER: US/08/776,188B  
 ; CURRENT FILING DATE: 2002-04-15  
 ; PRIOR APPLICATION NUMBER: PCT/EP95/02919  
 ; PRIOR FILING DATE: 1995-07-24  
 ; PRIOR APPLICATION NUMBER: P 44 30 972.4  
 ; PRIOR FILING DATE: 1994-08-31  
 ; PRIOR APPLICATION NUMBER: P 44 26 276.0  
 ; PRIOR FILING DATE: 1994-07-25  
 ; NUMBER OF SEQ ID NOS: 76  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 19  
 ; LENGTH: 19  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: SH-activated linear peptide 9  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (2)..(2)  
 ; OTHER INFORMATION: Xaa at position 2 is Beta-alanine  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (3)..(3)  
 ; OTHER INFORMATION: Xaa at position 3 is Epsilon-aminocaproic acid  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (4)..(4)  
 ; OTHER INFORMATION: Xaa at position 4 is Beta-alanine  
 ; US-08-776-188B-19

Query Match 41.6%; Score 72; DB 4; Length 19;  
 Best Local Similarity 66.7%; Pred. No. 0.00066;  
 Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 12 LNSMCKGRIICVYTS 26  
 1: |||||::|||  
 DB 5 LSLMCKGKGLVYTS 19

# RESULT 8

US-08-776-188B-36  
 ; Sequence 36, Application US/08776188B  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Weinhuess, Ursula-Henrike  
 ; APPLICANT: Kruse-Muller, Cornelia  
 ; APPLICANT: Hoss, Eva  
 ; APPLICANT: Faatz, Elke  
 ; APPLICANT: Offenloch-Hahnle, Beatus  
 ; APPLICANT: Seidel, Christoph  
 ; APPLICANT: Wiedmann, Michael  
 ; TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTIGEN  
 ; FILE REFERENCE: 100564-07003  
 ; CURRENT APPLICATION NUMBER: US/08/776,188B  
 ; CURRENT FILING DATE: 2002-04-15

;; PRIOR APPLICATION NUMBER: PCT/EP95/02919  
 ;; PRIOR FILING DATE: 1995-07-24  
 ;; PRIOR APPLICATION NUMBER: P 44 30 972.4  
 ;; PRIOR FILING DATE: 1994-08-31  
 ;; PRIOR APPLICATION NUMBER: P 44 26 276.0  
 ;; PRIOR FILING DATE: 1994-07-25  
 ;; NUMBER OF SEQ ID NOS: 76  
 ;; SOFTWARE: PatentIn version 3.1  
 ;; SEQ ID NO 36  
 ;; LENGTH: 19  
 ;; TYPE: PRT  
 ;; ORGANISM: Artificial Sequence  
 ;; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Digoxigenin-labelled line  
 ;; OTHER INFORMATION: gp41/2  
 ;; NAME/KEY: misc\_feature  
 ;; LOCATION: (1)..(1)  
 ;; OTHER INFORMATION: Xaa at position 1 is digoxigenin-3-cme  
 ;; FEATURE:  
 ;; NAME/KEY: misc\_feature  
 ;; LOCATION: (2)..(2)  
 ;; OTHER INFORMATION: Xaa at position 2 is Beta-alanine  
 ;; FEATURE:  
 ;; NAME/KEY: misc\_feature  
 ;; LOCATION: (3)..(3)  
 ;; OTHER INFORMATION: Xaa at position 3 is Epsilon-aminocaproic acid  
 ;; FEATURE:  
 ;; NAME/KEY: misc\_feature  
 ;; LOCATION: (4)..(4)  
 ;; OTHER INFORMATION: Xaa at position 4 is Beta-alanine  
 ;; US-08-776-188B-36

Query Match 41.6%; Score 72; DB 4; Length 19;  
 Best Local Similarity 66.7%; Pred. No. 0.00066;  
 Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 12 LNSMCKGRIICVYTS 26  
 1: |||||::|||  
 DB 5 LSLMCKGKGLVYTS 19

# RESULT 9

US-09-576-824A-20  
 ; Sequence 20, Application US/09576824A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: De Leys, Robert  
 ; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING  
 ; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN  
 ; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF  
 ; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORT  
 ; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS  
 ; TITLE OF INVENTION: CONTAINING THEM  
 ; FILE REFERENCE: 2752-11  
 ; CURRENT APPLICATION NUMBER: US/09/576,824A  
 ; CURRENT FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 08/723,425  
 ; PRIOR FILING DATE: 1996-09-30  
 ; PRIOR APPLICATION NUMBER: 09/146,028  
 ; PRIOR FILING DATE: 1993-11-22  
 ; PRIOR APPLICATION NUMBER: PCT/EP93/00517  
 ; PRIOR FILING DATE: 1993-03-08  
 ; PRIOR APPLICATION NUMBER: EP 92400596.6  
 ; PRIOR FILING DATE: 1992-03-06  
 ; NUMBER OF SEQ ID NOS: 600  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 20  
 ; LENGTH: 22  
 ; TYPE: PRT  
 ; ORGANISM: Human Immunodeficiency virus  
 ; FEATURE:  
 ; NAME/KEY: VARIANT

LOCATION: (1)  
OTHER INFORMATION: modified site  
NAME/KEY: VARIANT  
LOCATION: (22)  
OTHER INFORMATION: modified site  
US-09-576-824A-20

Query Match 41.6%; Score 72; DB 5; Length 22;  
Best Local Similarity 66.7%; Pred. No. 0.00077;  
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 MODQRLNSWCGCKRITC 23  
DB 4 LQDQRLNSWCGAFROYC 21

RESULT 10  
US-08-776-188B-10  
Sequence 10 Application US/08776188B  
GENERAL INFORMATION:  
APPLICANT: Weinues, Ursula-Henrike  
APPLICANT: Kruse-Muller, Cornelia  
APPLICANT: Hoss, Eva  
APPLICANT: Faatz, Elke  
APPLICANT: Ofenloch-Hahnle, Beatus  
APPLICANT: Seidel, Christoph  
APPLICANT: Wiedmann, Michael  
TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTIGEN  
FILE REFERENCE: 100564-07003  
CURRENT APPLICATION NUMBER: US/08/776,188B  
CURRENT FILING DATE: 2002-04-15  
PRIOR APPLICATION NUMBER: PCT/EP95/02919  
PRIOR FILING DATE: 1995-07-24  
PRIOR APPLICATION NUMBER: P 44 30 972.4  
PRIOR FILING DATE: 1994-08-31  
PRIOR APPLICATION NUMBER: P 44 26 276.0  
NUMBER OF SEQ ID NOS: 76  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 10  
LENGTH: 27  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Epitope region of HIV I or HI  
US-08-776-188B-10

Query Match 41.6%; Score 72; DB 4; Length 27;  
Best Local Similarity 52.4%; Pred. No. 0.00094;  
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 9 QOURLNSWCGKRIICYTSARW 29  
DB 1 QAOURLNSWCGAFROYCHTVPW 21

RESULT 11  
US-08-776-188B-54  
Sequence 54 Application US/08776188B  
GENERAL INFORMATION:  
APPLICANT: Weinues, Ursula-Henrike  
APPLICANT: Kruse-Muller, Cornelia  
APPLICANT: Hoss, Eva  
APPLICANT: Faatz, Elke  
APPLICANT: Ofenloch-Hahnle, Beatus  
APPLICANT: Seidel, Christoph  
APPLICANT: Wiedmann, Michael  
TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTIGEN  
FILE REFERENCE: 100564-07003  
CURRENT APPLICATION NUMBER: US/08/776,188B

CURRENT FILING DATE: 2002-04-15  
PRIOR APPLICATION NUMBER: PCT/EP95/02919  
PRIOR FILING DATE: 1995-07-24  
PRIOR APPLICATION NUMBER: P 44 30 972.4  
PRIOR FILING DATE: 1994-08-31  
PRIOR APPLICATION NUMBER: P 44 26 276.0  
NUMBER OF SEQ ID NOS: 76  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 54  
LENGTH: 31  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Ruthenylated linear peptide  
NAME/KEY: misc\_feature  
LOCATION: (1)..(1)  
OTHER INFORMATION: Xaa at position 1 is BPRu  
US-08-776-188B-54

Query Match 41.6%; Score 72; DB 4; Length 31;  
Best Local Similarity 52.4%; Pred. No. 0.0011;  
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 9 QOURLNSWCGKRIICYTSARW 29  
DB 5 QAOURLNSWCGAFROYCHTVPW 25

RESULT 12  
US-09-359-975-14  
Sequence 14 Application US/09359975  
GENERAL INFORMATION:  
APPLICANT: Weiner, David B.  
APPLICANT: Williams, William V.  
Mang, Bin  
TITLE OF INVENTION: Compositions and Methods for Delivery of  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz & Norris  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 mb-MD/JAF  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/359,975  
FILING DATE: 23-Jul-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,385B  
FILING DATE: 26-Nov-1997  
APPLICATION NUMBER: US 08/495,684  
FILING DATE: 28-Sep-1995  
APPLICATION NUMBER: PCT/US94/00899  
FILING DATE: 26-Jan-1994  
APPLICATION NUMBER: 08/125,012  
FILING DATE: 21-Sep-1993  
APPLICATION NUMBER: 08/124,962  
FILING DATE: 21-Sep-1993  
APPLICATION NUMBER: 08/093,235  
FILING DATE: 15-Jul-1993  
APPLICATION NUMBER: 08/029,336  
FILING DATE: 11-Mar-1993  
APPLICATION NUMBER: 08/008,342

FILING DATE: 26-JAN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Deluca, Mark  
 REGISTRATION NUMBER: 33,229  
 REFERENCE/DOCKET NUMBER: UPAD-0253  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-568-3100  
 TELEFAX: 215-568-3429  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 26 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
 US-09-359-975-14

Query Match 40.5%; Score 70; DB 5; Length 26;  
 Best Local Similarity 52.4%; Pred. No. 0.0018;  
 Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 3 ETLMDQRLNSWCKGRITC 23  
 Db 6 ERYIKDOLLGIWCGSKLIC 26

RESULT 13  
 US-08-776-188B-5  
 Sequence 5, Application US/08776188B  
 GENERAL INFORMATION:  
 APPLICANT: Weinhuess, Ursula-Henrike  
 APPLICANT: Kruse-Muller, Cornelia  
 APPLICANT: Hoss, Eva  
 APPLICANT: Paatz, Elke  
 APPLICANT: Offenloch-Hahnle, Beatus  
 APPLICANT: Seidel, Christoph  
 APPLICANT: Wiedmann, Michael  
 TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTIGEN  
 FILE REFERENCE: 100564-07003  
 CURRENT APPLICATION NUMBER: US/08/776,188B  
 CURRENT FILING DATE: 2002-04-15  
 PRIOR APPLICATION NUMBER: PCT/EP95/02919  
 PRIOR FILING DATE: 1995-07-24  
 PRIOR APPLICATION NUMBER: P 44 30 972.4  
 PRIOR FILING DATE: 1994-08-31  
 PRIOR APPLICATION NUMBER: P 44 26 276.0  
 PRIOR FILING DATE: 1994-07-25  
 NUMBER OF SEQ ID NOS: 76  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 5  
 LENGTH: 23  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Epitope region of HIV 1 or HI  
 OTHER INFORMATION: HIV subtype O  
 US-08-776-188B-5

Query Match 37.6%; Score 65; DB 4; Length 23;  
 Best Local Similarity 47.4%; Pred. No. 0.008;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 12 LNSWCKGRITCYTSARWH 30  
 Db 1 LGIWCGSKRLICTTAVPMW 19

RESULT 14  
 US-09-576-824A-6  
 Sequence 6, Application US/09576824A  
 GENERAL INFORMATION:

APPLICANT: De Leys, Robert  
 TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING  
 TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN  
 A PROCESS FOR DETERMINATION OF ANTIBODIES OF  
 BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORT  
 EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS  
 CONTAINING THEM  
 FILE REFERENCE: 2752-11  
 CURRENT APPLICATION NUMBER: US/09/576,824A  
 CURRENT FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 08/723,425  
 PRIOR FILING DATE: 1996-09-30  
 PRIOR APPLICATION NUMBER: 09/146,028  
 PRIOR FILING DATE: 1993-11-22  
 PRIOR APPLICATION NUMBER: PCT/EP93/00517  
 PRIOR FILING DATE: 1993-03-08  
 PRIOR APPLICATION NUMBER: EP 92400598.6  
 PRIOR FILING DATE: 1992-03-06  
 NUMBER OF SEQ ID NOS: 600  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 24  
 TYPE: PRT  
 ORGANISM: Human immunodeficiency virus  
 FEATURE:  
 NAME/KEY: VARIANT  
 LOCATION: (1)  
 OTHER INFORMATION: modified site  
 NAME/KEY: VARIANT  
 LOCATION: (24)  
 OTHER INFORMATION: modified site  
 US-09-576-824A-6

Query Match 32.1%; Score 55.5; DB 5; Length 24;  
 Best Local Similarity 47.8%; Pred. No. 0.19;  
 Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 8 DOORLNSWCKGRITCYTSARWH 30  
 Db 2 DOLLGIWCGSKRLIT-TVPMW 23

RESULT 15  
 US-08-776-188B-30  
 Sequence 30, Application US/08776188B  
 GENERAL INFORMATION:  
 APPLICANT: Weinhuess, Ursula-Henrike  
 APPLICANT: Kruse-Muller, Cornelia  
 APPLICANT: Hoss, Eva  
 APPLICANT: Paatz, Elke  
 APPLICANT: Offenloch-Hahnle, Beatus  
 APPLICANT: Seidel, Christoph  
 APPLICANT: Wiedmann, Michael  
 TITLE OF INVENTION: DETERMINATION OF A SPECIFIC IMMUNOGLOBULIN USING MULTIPLE ANTI  
 FILE REFERENCE: 100564-07003  
 CURRENT APPLICATION NUMBER: US/08/776,188B  
 CURRENT FILING DATE: 2002-04-15  
 PRIOR APPLICATION NUMBER: PCT/EP95/02919  
 PRIOR FILING DATE: 1995-07-24  
 PRIOR APPLICATION NUMBER: P 44 30 972.4  
 PRIOR FILING DATE: 1994-08-31  
 PRIOR APPLICATION NUMBER: P 44 26 276.0  
 PRIOR FILING DATE: 1994-07-25  
 NUMBER OF SEQ ID NOS: 76  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 30  
 LENGTH: 20  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Digoxigenin-labelled linea  
 OTHER INFORMATION: gp41/2

```

: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(1)
: OTHER INFORMATION: Xaa at position 1 is digoxigenin-3-cme
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2)..(2)
: OTHER INFORMATION: Xaa at position 2 is Beta-alanine
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (3)..(3)
: OTHER INFORMATION: Xaa at position 3 is Epsilon-aminocaproic acid
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (4)..(4)
: * OTHER INFORMATION: Xaa at position 4 is Beta-alanine
US-08-776-188B-30

```

```

Query Match      31.8%; Score 55; DB 4; Length 20;
Best Local Similarity 53.3%; Pred. No. 0.19;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 12 LNSWCKKGRITCYTS 26
   | || | : || | :
Db 5 LGIMCCSGKLICTTA 19

```

Search completed: June 20, 2002, 15:34:57  
 Job time: 205 sec





C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S21990; S70423  
R:Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined  
A:Reference number: S21990  
A:Accession: S21990  
A:Molecule type: DNA  
A:Residues: 1-357 <STE1>  
A:Cross-references: EMBL:X61357; NID:g60175; PID:CAA43626.1; PID:g60176  
R:Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid  
A:Reference number: S70417; MUID:92144209  
A:Accession: S70423  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-332, 'X', 334-357 <STE2>  
A:Cross-references: EMBL:X61357; NID:g60175; PID:CAA43626.1; PID:g60176  
C:Superfamily: type E retrovirus env polyprotein

Query Match 50.9%; Score 88; DB 2; Length 357;  
Best Local Similarity 50.0%; Pred. No. 7.2e-05;  
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWCKGRITCYTSARWH 30  
| : : | | | | | | | | | | : : |  
DB 85 ERYLKDQDLGLTIGWCGSRILCTTAVPMW 113

RESULT 8  
S22002  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 3L  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S22002; S70418  
R:Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined  
A:Reference number: S21990  
A:Accession: S22002  
A:Molecule type: DNA  
A:Residues: 1-358 <STE1>  
A:Cross-references: EMBL:X61352; NID:g60186; PID:CAA43616.1; PID:g60187  
R:Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid  
A:Reference number: S70417; MUID:92144209  
A:Accession: S70418  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333, 'X', 335-358 <STE2>  
A:Cross-references: EMBL:X61352; NID:g60186  
C:Superfamily: type E retrovirus env polyprotein

Query Match 50.9%; Score 88; DB 2; Length 358;  
Best Local Similarity 50.0%; Pred. No. 7.2e-05;  
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWCKGRITCYTSARWH 30  
| : : | | | | | | | | | | : : |  
DB 86 ERYLKDQDLGLTIGWCGSRILCTTAVPMW 113

RESULT 9  
S22000  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 12-Apr-1995  
C:Accession: S22000

R:Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determined  
A:Reference number: S21990  
A:Accession: S22000  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <STE>  
A:Cross-references: EMBL:X61351  
C:Superfamily: type E retrovirus env polyprotein

Query Match 50.3%; Score 87; DB 2; Length 358;  
Best Local Similarity 50.0%; Pred. No. 0.0001;  
Matches 14; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWCKGRITCYTSARWH 30  
| : : | | | | | | | | | | : : |  
DB 86 ERYLKDQDLGLTIGWCGSRILCTTAVPMW 113

RESULT 10  
S70417  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 3b) (fragment)  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: patient 3B  
C>Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 26-Aug-1999  
C:Accession: S70417  
R:Steuier, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid  
A:Reference number: S70417; MUID:92144209  
A:Accession: S70417  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <STE>  
A:Cross-references: EMBL:X61351; NID:g60184; PID:CAA43614.1; PID:g60185  
C:Superfamily: type E retrovirus env polyprotein

Query Match 50.3%; Score 87; DB 2; Length 358;  
Best Local Similarity 50.0%; Pred. No. 0.0001;  
Matches 14; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWCKGRITCYTSARWH 30  
| : : | | | | | | | | | | : : |  
DB 86 ERYLKDQDLGLTIGWCGSRILCTTAVPMW 113

RESULT 11  
VCLUST  
env polyprotein precursor - human immunodeficiency virus type 2 (isolate SF)  
N:Alternate names: coat polyprotein  
N:Contains: surface glycoprotein gp120; transmembrane glycoprotein gp41  
C:Species: human immunodeficiency virus type 2, HIV-2  
A:Note: host Homo sapiens (man)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 09-Sep-1994  
C:Accession: H33943  
R:Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.  
J. Virol. 64, 890-901, 1990  
A:Title: Molecular characterization of an attenuated human immunodeficiency virus type 2  
A:Reference number: A33943; MUID:90112662  
A:Accession: H33943  
A:Molecule type: genomic RNA  
A:Residues: 1-859 <KUM>  
A:Cross-references: EMBL:M86924  
C:Genetics: env  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-859/Product: env polyprotein #status predicted <ENV>  
F:20-501/Product: surface glycoprotein gp120 #status predicted <SGG>



Query Match 49.7%; Score 86; DB 2; Length 151;  
 Best Local Similarity 48.1%; Pred. No. 6.2e-05;  
 Matches 13; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 3 ETLMDQDQRLNSWCGCKGRICYSARW 29  
 |::||:||||| |::| |  
 DB 28 EKYLNDAQLNSWGCAPROVCHTVEW 54

Search completed: June 20, 2002, 15:32:12  
 Job time: 220 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 15:31:52 ; Search time 10.33 Seconds  
(without alignments) 112.448 Million cell updates/sec

Title: US-09-605-573a-69  
Perfect score: 173  
Sequence: 1 GRETLMQDQRLNSMCKGRICYSARWH 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSPROT\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	93	53.8	ENV_SIVCZ	P17281 chimpanzee
2	93	53.8	ENV_HV2KR	Q74126 human immun
3	92	53.2	ENV_HV2RO	P04577 human immun
4	89	51.4	ENV_SIVAT	P05886 simian immu
5	87	50.3	ENV_SIVAT	Q02837 simian immu
6	87	50.3	ENV_HV2NZ	P05883 human immun
7	86	49.7	ENV_HV2SB	P12449 human immun
8	86	49.7	ENV_HV1ZH	P05881 human immun
9	85	49.1	ENV_HV1Y2	P35961 human immun
10	85	49.1	ENV_HV1S1	P19550 human immun
11	85	49.1	ENV_HV1W2	P04582 human immun
12	85	49.1	ENV_HV1B8	P14488 human immun
13	85	49.1	ENV_HV1BN	P19549 human immun
14	85	49.1	ENV_HV1S3	P19551 human immun
15	85	49.1	ENV_HV1MF	P12487 human immun
16	85	49.1	ENV_HV1Z2	P03378 human immun
17	85	49.1	ENV_HV1A2	P03888 human immun
18	85	49.1	ENV_HV1OY	P04580 human immun
19	85	49.1	ENV_HV1Z6	P03375 human immun
20	85	49.1	ENV_HV1B1	Q70626 human immun
21	85	49.1	ENV_HV1H2	P04579 human immun
22	85	49.1	ENV_HV1LW	P05877 human immun
23	85	49.1	ENV_HV1LM	P03376 human immun
24	85	49.1	ENV_HV1PY	P03878 human immun
25	85	49.1	ENV_HV1SC	P03878 human immun
26	85	49.1	ENV_HV1WL	P03878 human immun
27	85	49.1	ENV_HV1BR	P03377 human immun
28	85	49.1	ENV_HV1RH	P04579 human immun
29	85	49.1	ENV_HV1J3	P12489 human immun
30	85	49.1	ENV_HV1C4	P05879 human immun
31	84	48.6	ENV_HV2S2	P32536 human immun
32	84	48.6	ENV_HV2D1	P17755 human immun
33	84	48.6	ENV_HV2G1	P18040 human immun

34	84	48.6	859	1	ENV_HV2CA	P24105 human immun
35	84	48.6	859	1	ENV_HV2D2	P15831 human immun
36	84	48.6	859	1	ENV_HV2ST	P20872 human immun
37	84	48.6	877	1	ENV_SIVAG	P27977 simian immu
38	84	48.6	885	1	ENV_SIVS4	P12492 simian immu
39	83	48.0	846	1	ENV_HV1ND	P16799 human immun
40	83	48.0	848	1	ENV_HV1JR	P20871 human immun
41	83	48.0	856	1	ENV_HV1H3	P04624 human immun
42	82	47.4	863	1	ENV_HV1Z8	P05882 human immun
43	82	47.4	881	1	ENV_SIVMK	P05884 simian immu
44	81	46.8	380	1	ENV_SIVM2	P08810 simian immu
45	81	46.8	768	1	ENV_SIVV1	P27757 simian immu

## ALIGNMENTS

RESULT ID	ENV_SIVCZ	STANDARD:	PRT:	854 AA.
AC	P17281:			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11723;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90259077; PubMed=2188136;			
RA	Huet T., Cheynier R., Meyers A., Roelants G., Wain-Hobson S.;			
RT	"Genetic organization of a chimpanzee lentivirus related to HIV-1."			
RL	Nature 345:356-359(1990).			
CC	- I - SIMILARITY: STRONG, WITH HIV-1 ENV POLYPEPTIDE.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sb.ch/announce/">http://www.isb-sb.ch/announce/</a> or send an email to <a href="mailto:license@isb-sb.ch">license@isb-sb.ch</a> ).			
CC	-----			
DR	EMBL: X52154; CA36407.1; -			
DR	PIR: S09990; VCLJST.			
DR	HIV: X52154; ENVSCP2.			
DR	InterPro: IPR000328; ENV_GP41.			
DR	InterPro: IPR000777; GP120.			
DR	Pfam: PF00516; GP120; 1.			
DR	Pfam: PF00517; GP41; 1.			
KW	AIDS; Coat protein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	500	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	501	854	TRANSMEMBRANE GLYCOPROTEIN.
FT	TRANSMEM	501	517	POTENTIAL.
FT	TRANSMEM	675	693	POTENTIAL.
FT	TRANSMEM	805	821	POTENTIAL.
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	299	299	N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	331	331	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	336	336	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	351	351	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	356	356	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	384	384	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	392	392	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	426	426	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	432	432	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	446	446	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	450	450	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	601	601	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	608	608	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	616	616	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	628	628	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
SO	SEQUENCE	854 AA;	95803 MM;	2E249AFAD4F2D9B3	CRC64;			
Query Match								
Best Local Similarity		53.8%;	Score 93;	DB 1;	Length 854;			
Matches 14;		Conservative	5;	Mismatches	9;	Indels 0;		
					Gaps	0;		
Oy	3 ETLMODQRLNSGCKGRICYSARWH 30							
	: : : : :   : : : : :   : : : : :   : : : : :   : : : : :   : : : : :							
Db	574 ERYLDQDQILGMGCGKAVCTTYPWN 601							
RESULT	2							
ENV_HV2KR								
ID	ENV_HV2KR	STANDARD;	PRT;	857 AA.				
AC	074126;							
DT	01-NOV-1997 (Rel. 35, Created)							
DT	01-NOV-1997 (Rel. 35, Last sequence update)							
DT	15-JUL-1999 (Rel. 38, Last annotation update)							
DE	Envelope polypeptide gp160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].							
GN	ENV.							
OS	Human immunodeficiency virus type 2 (isolate KR) (HIV-2).							
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.							
OK	NCBI_TaxID=73484;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RA	Kraus G.K., Talbot R., Leavitt M., Luzznick L., Schmidt A., Babel P., Bartz C., Morton W., Wong-Staal F., Looney D.J.;							
RL	Submitted (Apr-1995) to the EMBL/Genbank/DBJ databases.							
CC	-----							
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on ways to use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).							
CC	-----							
DR	EMBL; U22047; AAA64582.1; -							
DR	InterPro: IPR000328; Env_GP41.							
DR	InterPro: IPR000777; GP120.							
DR	Pfam: PF00516; GP120.1.							
KM	Pfam: PF00517; GP41.1.							
DR	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.							
KW	Signal.							
FT	SIGNAL	1	19	POTENTIAL.				
FT	CHAIN	20	503	EXTERIOR MEMBRANE GLYCOPROTEIN.				
FT	CHAIN	504	857	TRANSMEMBRANE GLYCOPROTEIN.				
FT	CARBOHYD	36	36	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	69	69	N-LINKED	(GLCNAC. . .)	(POTENTIAL)		
FT	CARBOHYD	78	78	N-L				

FT	CARBOHYD	230	230	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	264	264	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	281	281	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	302	302	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	358	358	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	391	391	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	440	440	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	455	455	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	612	612	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
FT	CARBOHYD	628	628	N-LINKED (GLCNAC. . .)	(POTENTIAL).		
SO	SEQUENCE	857 AA;	98689 MW;	P881C6755B5746DF	CNC64;		
Query Match		53.8%;	Score 93;	DB 1;	Length 857;		
Best Local Similarity		55.6%;	Pred. No. 2.7e-06;				
Matches 15;		Conservative 4;	No. matches 8;	Indels 0;	Gaps 0;		
Oy	3	ETLMDDQRLNSGCKGRLLICYSRW	29				
...							
Db	575	EKYLDDQARLNSGCAFRQVCITYLW	601				
RESULT	3						
ENV_HV2RO	ENV_HV2RO	STANDARD;	PRT;	858 AA.			
AC	P04577;						
DT	13-AUG-1987 (Rel. 05, Created)						
DT	13-AUG-1987 (Rel. 05, Last sequence update)						
DT	16-OCT-2001 (Rel. 40, Last annotation update)						
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].						
DE	ENV.						
GN	Human immunodeficiency virus type 2 (isolate ROD) (HIV-2).						
OS	Viruses; Retroviral viruses; Retroviridae; Lentivirus.						
OC	NCBI_TaxID=11720;						
OX	[1]						
RN	SEQUENCE FROM N.A.						
RP	MEDLINE=87173056; PubMed=3031510;						
RX	Guyader M., Emerman M., Sonigo P., Clavel F., Montagnier L.,						
RA	Allison M.;						
RT	"Genome organization and transactivation of the human						
RT	immunodeficiency virus type 2.";						
RL	Nature 326:662-669(1987).						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						
CC	use by non-profit institutions as long as its content is in no way						
CC	modified and this statement is not removed. Usage by and for commercial						
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).						
CC	-----						
DR	EMBL; M15390; AAB00770.1; -.						
DR	EMBL; X05291; CAA28914.1; -.						
DR	PIR; C26262; VCL02.						
DR	HIV; M15390; ENV52ROD.						
DR	InterPro; IPR000328; ENV_GP41.						
DR	InterPro; IPR000777; GP120.						
DR	Pfam; PF00516; GP120.1.						
DR	Pfam; PF00517; GP41.1.						
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;						
KW	Signal.						
FT	SIGNAL	1	17				
FT	CHAIN	18	501				
FT	CHAIN	502	858				
FT							

```

FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 312 312 T -> I (IN REF 1; AAB00770).
SQ SEQUENCE 838 AA; 98824 MW; C7266AF1F5C5B9A7 CRC64;

```

```

Query Match 53.2%; Score 92; DB 1; Length 858;
Best Local Similarity 55.6%; Pred. No. 3.9e-06;
Matches 15; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 3 ETLMDQORLNSMGCGRIICYSARW 29
DB 583 EKYLDQARLNSMGCAFRCVCHTVPW 609

```

```

RESULT 4
ENV_SIVAT STANDARD: PRT; 865 AA.
ID ENV_SIVAT
AC P05886;
DT 01-NOV-1988 (Rel. 09; Last sequence update)
DT 01-NOV-1988 (Rel. 40; Last annotation update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Simian immunodeficiency virus (Tyo-1 isolate) (SIV-AGM).
OC Viruses: Retrovirda; Retroviridae; Lentiviridae.
OX NCBI_TaxID=11731;
RN [1]
RP MEDLINE=88232906; PubMed=3374586;
RA Fukusawa M., Mura T., Hasegawa A., Morikawa S., Tsujimoto H.,
RT "Sequence of simian immunodeficiency virus from African green monkey,
RT a new member of the HIV/SIV group.";
RL Nature 333:457-461(1988).
CC -1- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X07805; CAA30663.2; -
CC PIR; G30045; VCLG4.

```

```

DR HIV; X07805; ENV5AGMTV.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Aids; Coat protein; Glycoprotein; Polypeptide; Transmembrane;
KW Signal.
FT CHAIN 1 20 POTENTIAL.
FT CHAIN 21 536 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 537 865 TRANSMEMBRANE GLYCOPROTEIN.
FT SITE 770 770 IN-FRAME TERMINATION CODON.
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 645 645 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 99024 MW; 6CEFOF09001D6D95 CRC64;

```

```

Query Match 51.4%; Score 89; DB 1; Length 865;
Best Local Similarity 48.1%; Pred. No. 1.1e-05;
Matches 13; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 3 ETLMDQORLNSMGCGRIICYSARW 29
DB 608 EKYLDQARLNSMGCAFRCVCHTVPW 634

```

```

RESULT 5
ENV_SIVAT STANDARD: PRT; 854 AA.
ID ENV_SIVAT
AC 002837;
DT 01-JUL-1993 (Rel. 26; Last sequence update)
DT 01-JUL-1993 (Rel. 26; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Simian immunodeficiency virus (isolate AGM / clone GRI-1) (SIV-AGM).
OC Viruses: Retrovirda; Retroviridae; Lentiviridae.
OX NCBI_TaxID=31684;
RN [1]
RP MEDLINE=91220680; PubMed=2024476;
RA Fomsgaard A., Hirsch V.M., Allan J.S., Johnson P.R.;
RT "A highly divergent proviral DNA clone of SIV from a distinct species
RT of African green monkey.";
RL Virology 182:397-402(1991).
CC -1- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC -----
CC CC

```

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M66437; AAA91928.1; -  
DR EMBL; M58410; AAA47591.1; -  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 522 EXTERIOR MEMBRANE GLYCOPROTEIN  
FT CHAIN 523 854 TRANSMEMBRANE GLYCOPROTEIN  
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 854 AA; 96855 MW; 5919CA6C9622912F CRC64;

Query Match 50.3%; Score 87; DB 1; Length 854;  
Best Local Similarity 48.1%; Pred. No. 2.2e-05;  
Matches 13; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 3 ETLMDQORLNSGCKGRITCYTSARW 29  
| : : : | : : : | : : : | : : : | : : : |  
Db 594 EKYLEDQARLNSGCKAMKQVCHTTPW 620

RESULT 6  
ENV\_HV2NZ STANDARD; PRT; 856 AA.  
AC P05883;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 2 (isolate NIH-2) (HIV-2).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11719;  
RN [1]  
SEQUENCE FROM N.A.  
RX MEDLINE=88320359; PubMed=3261862;  
RA Zagury J.F., Francini G., Reitz M.S. Jr., Collalti E., Starcich B.R., Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laure F., Ayra S.K., Josephs S.F., Zagury D., Wong-Staal F., Gallo R.C.;  
RT Genetic variability between isolates of human immunodeficiency virus (HIV) type 2 is comparable to the variability among HIV type 1."

RL Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; J03654; AAB00761.1; -  
DR HIV; J03654; ENV52NTHZ.  
DR InterPro; IPR000328; Env\_GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 492 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 493 856 TRANSMEMBRANE GLYCOPROTEIN.  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 96655 MW; A938B0A7E2B881D6 CRC64;

Query Match 50.3%; Score 87; DB 1; Length 856;  
Best Local Similarity 51.9%; Pred. No. 2.2e-05;  
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 ETLMDQORLNSGCKGRITCYTSARW 29  
| : : : | : : : | : : : | : : : | : : : |  
Db 574 EKYLKDAQLNSGCKAFQVCHTSPW 600

RESULT 7  
ENV\_HV2SB STANDARD; PRT; 846 AA.  
AC P12449;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 2 (isolate SBLISY) (HIV-2).  
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11718;

[illegible]



FT DISULFID 125 194 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 216 245 BY SIMILARITY.  
 FT DISULFID 226 237 BY SIMILARITY.  
 FT DISULFID 294 328 BY SIMILARITY.  
 FT DISULFID 374 435 BY SIMILARITY.  
 FT DISULFID 381 408 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7F2AB CRC64;

Query Match 49.1%; Score 85; DB 1; Length 847;  
 Best Local Similarity 46.4%; Pred. No. 4.4e-05;  
 Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Oy 3 ETLMDQRLNSWGCKGRICYSARWH 30  
 Db 575 ERYLDDQLGLWGSGKLCITTVPMW 602

RESULT 11  
 ENV\_HV1W2 STANDARD; PRT; 847 AA.  
 AC P05880;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV  
 OS Human immunodeficiency virus type 1 (WU2 isolate) (HIV-1).  
 OC Viruses; Retrovird viruses; Retroviridae; Lentiviridae.  
 CC NCBI\_TaxID=11705;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86235450; PubMed=3012778;  
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D., Sabinuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;  
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS."  
 RL Science 232:1548-1553(1986).  
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-stb.ch/announce/> or send an email to [license@isb-stb.ch](mailto:license@isb-stb.ch)).  
 CC -----

DR EMBL: M12507; AAB12990.1; -  
 DR HIV: M12507; ENV5WMD2.  
 DR InterPro: IPR000328; ENV\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DR Pfam: PF00517; GP41.1.  
 KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane;

KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 501  
 FT CHAIN 502 847  
 FT DISULFID 53 73 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 118 202 BY SIMILARITY.  
 FT DISULFID 125 193 BY SIMILARITY.  
 FT DISULFID 130 152 BY SIMILARITY.  
 FT DISULFID 215 244 BY SIMILARITY.  
 FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 293 326 BY SIMILARITY.  
 FT DISULFID 372 435 BY SIMILARITY.  
 FT CARBOHYD 379 408 BY SIMILARITY.  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 847 AA; 96466 MW; CD1E33D73AASBCAE CRC64;

Query Match 49.1%; Score 85; DB 1; Length 847;  
 Best Local Similarity 46.4%; Pred. No. 4.4e-05;  
 Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Oy 3 ETLMDQRLNSWGCKGRICYSARWH 30  
 Db 575 ERYLDDQLGLWGSGKLCITTVPMW 602

RESULT 12  
 ENV\_HV1B8 STANDARD; PRT; 851 AA.  
 AC P04582;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV\_HV1B8

[illegible]

SO	SEQUENCE	851 AA;	96644 MW;	D16A3C90857785F1	CRC64;
	Query Match	49.18;	Score 85;	DB 1;	Length 851;
	Best Local Similarity	46.48;	Pred. No. 4,4e-05;		
	Matches 13;	Conservative 6;	Mismatches 9;	Indels 0;	Gaps 0;
OY	3	ETLMDOORLNSMGKGRITCYTSARWH	30		
	1			1	1
Db	579	ERYLKDQLLGICMGCKRLICTTAVPWN	606		
	RESULT 13				
ENV_HV1BN	ENV_HV1BN	STANDARD;	PRT;	852 AA.	
AC	PI2488;				
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-OCT-1989 (Rel. 12, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)]				
DE	ENV.				
OS	Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).				
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID=11693;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89085613; PubMed=2789516;				
RA	Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P., Dandekar S.,				
RT	"Biological and molecular characterization of human immunodeficiency virus (HIV-1BR) from the brain of a patient with progressive dementia."				
RL	Virology 168:79-89(1989).				
CC	-I- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL, M21098; AAA44221.1; -				
DR	PIR; A31667; VCLJBR.				
DR	HIV; M21098; ENVSBRVA.				
DR	InterPro; IPR000328; Env_GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
KW	AIDS; Coat protein; Polypeptide; Polypeptide; Transmembrane; Signal.				
FT	SIGNAL	1	30		
FT	CHAIN	31	507		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	508	852		TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74		BY SIMILARITY.
FT	DISULFID	119	205		BY SIMILARITY.
FT	DISULFID	126	196		BY SIMILARITY.
FT	DISULFID	131	155		BY SIMILARITY.
FT	DISULFID	218	247		BY SIMILARITY.
FT	DISULFID	228	239		BY SIMILARITY.
FT	DISULFID	296	330		BY SIMILARITY.
FT	DISULFID	376	439		BY SIMILARITY.
FT	DISULFID	383	412		BY SIMILARITY.
FT	CARBOHYD	49	49		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	88	88		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	135	135		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	138	138		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	157	197		N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 852 AA; 97203 MW; 2BB86345DEC915F CRC64;

Query Match 49.1%; Score 85; DB 1; Length 852;  
 Best Local Similarity 46.4%; Pred. No. 4.4e-05;  
 Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 3 ETLMDQDRLNSMCGKRICYTSARWH 30  
 Db 580 ERYLKDQQLGIMGCSKLCITTVPMW 607

RESULT 14  
 ID ENV\_HV1S3 STANDARD: PRT: 852 AA.  
 AC P19549:  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317906; PubMed=2370688;  
 RA York-Higgins D., Cheng-Wayer C., Bauer D., Levy J.A., Dina D.;  
 RT "Human immunodeficiency virus type 1 cellular host range, replication, and cytopathicity are linked to the envelope region of the viral genome.";  
 RL J. Virol. 64:4016-4020(1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M38427; AAA45067.1; -  
 DR HIV: M38427; ENVSEF3.  
 DR InterPro: IPR000328; ENV\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.  
 KW Signal.  
 FT SIGNAL 1 31 BY SIMILARITY.

FT CHAIN 32 506  
 FT CHAIN 507 852  
 FT DISULFID 53 73  
 FT DISULFID 118 125  
 FT DISULFID 125 197  
 FT DISULFID 130 156  
 FT DISULFID 219 248  
 FT DISULFID 229 240  
 FT DISULFID 297 331  
 FT DISULFID 377 439  
 FT DISULFID 384 412  
 FT CARBOHYD 87 87  
 FT CARBOHYD 129 129  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 142 142  
 FT CARBOHYD 155 155  
 FT CARBOHYD 159 159  
 FT CARBOHYD 189 189  
 FT CARBOHYD 198 198  
 FT CARBOHYD 242 242  
 FT CARBOHYD 263 263  
 FT CARBOHYD 277 277  
 FT CARBOHYD 290 290  
 FT CARBOHYD 296 296  
 FT CARBOHYD 332 332  
 FT CARBOHYD 339 339  
 FT CARBOHYD 355 355  
 FT CARBOHYD 385 385  
 FT CARBOHYD 391 391  
 FT CARBOHYD 397 397  
 FT CARBOHYD 401 401  
 FT CARBOHYD 405 405  
 FT CARBOHYD 442 442  
 FT CARBOHYD 457 457  
 FT CARBOHYD 607 607  
 FT CARBOHYD 612 612  
 FT CARBOHYD 621 621  
 FT CARBOHYD 633 633  
 FT CARBOHYD 812 812  
 SQ SEQUENCE 852 AA; 96663 MW; E27BFB823C9910D CRC64;

Query Match 49.1%; Score 85; DB 1; Length 852;  
 Best Local Similarity 46.4%; Pred. No. 4.4e-05;  
 Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 3 ETLMDQDRLNSMCGKRICYTSARWH 30  
 Db 580 ERYLKDQQLGIMGCSKLCITTVPMW 607

RESULT 15  
 ID ENV\_HV1MF STANDARD: PRT: 853 AA.  
 AC P19551:  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11704;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317877; PubMed=1695254;  
 RA Stevenson M., Haggerty S., Lamontica C., Mann A.M., Meter C., Wasilak A.;  
 RT "Cloning and characterization of human immunodeficiency virus type 1 variants diminished in the ability to induce syncytium-independent cytolysis.";  
 RT



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 15:31:02 ; Search time 25.12 seconds  
(without alignments)  
206,602 Million cell updates/sec

Title: US-09-605-573A-69  
Perfect score: 173  
Sequence: 1 GRETLMODQORLNSWCKGRICYSARWH 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: SP\_archaea:\*  
2: SP\_bacteria:\*  
3: SP\_fungi:\*  
4: SP\_human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133	76.9	161	15	Q9IEB5 human immun
2	133	76.9	524	15	Q9IEB3 human immun
3	132	76.3	179	15	Q9IEB4 human immun
4	132	76.3	195	15	Q9IEC6 human immun
5	132	76.3	576	15	Q9IEF1 human immun
6	131	75.7	219	15	Q9IEC8 human immun
7	130	75.1	114	15	Q9IEC8 human immun
8	129	74.6	216	15	Q9IEC5 human immun
9	129	74.6	219	15	Q9IEB6 human immun
10	128	74.0	120	15	Q9IEB6 human immun
11	128	74.0	134	15	Q9IEB6 human immun
12	128	74.0	143	15	Q9IEB6 human immun
13	128	74.0	143	15	Q9IEB6 human immun
14	128	74.0	143	15	Q9IEB6 human immun
15	127	73.4	105	15	Q9IEB6 human immun
16	127	73.4	172	15	Q9IEB3 human immun

17	127	73.4	234	15	Q9IEC2 human immun
18	127	73.4	242	15	Q9IE31 human immun
19	126	72.8	125	15	Q9IEH8 human immun
20	126	72.8	130	15	Q9IEH9 human immun
21	126	72.8	216	15	Q9IEA5 human immun
22	126	72.8	232	15	Q9IEB1 human immun
23	126	72.8	242	15	Q9IE30 human immun
24	126	72.8	243	15	Q9IE55 human immun
25	126	72.8	544	15	Q9IEB9 human immun
26	125	72.3	111	15	Q9IEB9 human immun
27	125	72.3	114	15	Q9IEB9 human immun
28	125	72.3	114	15	Q9IEB9 human immun
29	125	72.3	224	15	Q9IEA8 human immun
30	125	72.3	418	15	Q9IEA7 human immun
31	124	71.7	124	15	Q9IEH7 human immun
32	124	71.7	126	15	Q9IEH1 human immun
33	124	71.7	134	15	Q9IEH4 human immun
34	124	71.7	183	15	Q9IEC1 human immun
35	124	71.7	200	15	Q9IEB8 human immun
36	124	71.7	213	15	Q9IEC4 human immun
37	124	71.7	220	15	Q9IEC9 human immun
38	124	71.7	225	15	Q9IEA0 human immun
39	124	71.7	512	15	Q9IEB2 human immun
40	124	71.7	517	15	Q9IEE7 human immun
41	124	71.7	532	15	Q9IEF0 human immun
42	124	71.7	545	15	Q9IEB7 human immun
43	123	71.1	112	15	Q9IEB7 human immun
44	123	71.1	114	15	Q9IEB7 human immun
45	123	71.1	116	15	Q9IEB7 human immun

## ALIGNMENTS

RESULT 1  
ID Q9IEB5 PRELIMINARY: PRT: 161 AA.  
AC Q9IEB5;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE GP41 (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCF57;  
RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
RT Philippe M.;  
\*Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates.\*;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ236405; CAB96253.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 161  
SQ SEQUENCE 161 AA; 19305 MW; 5E3AF197E1FDEB7C CRC64;

Query Match 76.9%; Score 133; DB 15; Length 161;  
Best Local Similarity 78.6%; Pred. No. 7.5e-13;  
Matches 22; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWCKGRICYSARWH 30  
DB 31 ETLMDQORLNSWCKGRICYSARWH 30  
RESULT 2  
Q9IEB3

ID Q9IED3 PRELIMINARY; PRT; 524 AA.  
 AC Q9IED3:  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV POLYPEPTIDE (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF57;  
 RA Roques P., Robertson D., Diamond F., Souquiere S., Mauchere P.,  
 RA Deplenne C., Brun-Vezinet F., Dormont D.;  
 RT "HIV-1 group O phylogenetic analysis of C2-gp41 region."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ133074; CAB96235.1; -  
 DR EMBL; AJ236406; CAB96242.1; -  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam; PF00517; GP41; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 FT NON\_TER 1 524  
 FT NON\_TER 1 524  
 SQ SEQUENCE 524 AA; 59109 MW; C3D9FF12207AEB41 CRC64;

Query Match 76.3%; Score 132; DB 15; Length 524;  
 Best Local Similarity 78.6%; Pred. No. 2.7e-12;  
 Matches 22; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 355 ETLIONQORLNSWGCKGRITCYTSARWH 382

RESULT 3  
 O9IEB4 PRELIMINARY; PRT; 179 AA.  
 AC O9IEB4:  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Gp41 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF58;  
 RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ236406; CAB96254.1; -  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam; PF00517; GP41; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR Transmembrane.  
 FT NON\_TER 1 179  
 FT NON\_TER 1 179  
 SQ SEQUENCE 179 AA; 21225 MW; B060634FF1D4531B CRC64;

Query Match 76.3%; Score 132; DB 15; Length 179;  
 Best Local Similarity 78.6%; Pred. No. 1.2e-12;  
 Matches 22; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 34 ETLIONQORLNSWGCKGRITCYTSARWH 61

RESULT 4  
 O9IEC6 PRELIMINARY; PRT; 195 AA.  
 AC O9IEC6:  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE GP41 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF06;  
 RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ236393; CAB96242.1; -  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR Transmembrane.  
 FT NON\_TER 1 195  
 FT NON\_TER 1 195  
 SQ SEQUENCE 195 AA; 22964 MW; 2A63633CD0C86A85 CRC64;

Query Match 76.3%; Score 132; DB 15; Length 195;  
 Best Local Similarity 78.6%; Pred. No. 1.3e-12;  
 Matches 22; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 35 ETLIONQORLNSWGCKGRITCYTSARWH 62

RESULT 5  
 O9IEF1 PRELIMINARY; PRT; 576 AA.  
 AC O9IEF1:  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV POLYPEPTIDE (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF06;  
 RA Roques P., Robertson D., Diamond F., Souquiere S., Mauchere P.,  
 RA Deplenne C., Brun-Vezinet F., Dormont D.;  
 RT "HIV-1 group O phylogenetic analysis of C2-gp41 region."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ133056; CAB96217.1; -  
 DR EMBL; AJ236406; CAB96217.1; -  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam; PF00517; GP41; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 FT NON\_TER 1 576  
 FT NON\_TER 1 576  
 SQ SEQUENCE 576 AA; 64977 MW; A21509F3C618195D CRC64;

Query Match 76.3%; Score 132; DB 15; Length 576;  
 Best Local Similarity 78.6%; Pred. No. 4.3e-12;  
 Matches 22; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSWGCKGRITCYTSARWH 30

DB 392 ETLIONOOLNSWGCKGRICVTSYKWN 419

## RESULT 6

AC 091EC8 PRELIMINARY; PRT; 219 AA.  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE GP41 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF02;  
 RA - Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL EMBL: AJ236391: CAB96240.1; to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ236391: CAB96240.1; to the EMBL/GenBank/DBJ databases.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam: PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 219 219  
 SQ SEQUENCE 219 AA; 25363 MW; 85C2DC5B5F528907 CRC64;

## Query Match

Best Local Similarity 75.7%; Score 131; DB 15; Length 219;  
 Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ETLMOOQRLNSWGCKGRICVTSYKWN 30  
 DB 48 ETLIONOOLNSWGCKGRICVTSYKWN 75

## RESULT 7

ID 040456 PRELIMINARY; PRT; 114 AA.  
 AC 040456;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ENVELOPE TRANSMEMBRANE GLYCOPROTEIN (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GROUP O;  
 RA Bibollet-Ruche F., Ekasa E., Peeters M., Delaporte E.;  
 RT "Molecular characterization of envelope transmembrane glycoprotein of  
 RT 14 new human immunodeficiency virus type 1 group O strains from  
 RT different African countries."  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y09772: CAAT0911.1; to the EMBL/GenBank/DBJ databases.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam: PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 114 114  
 SQ SEQUENCE 114 AA; 13737 MW; 7907E77F5B51FA9D CRC64;

Query Match 75.1%; Score 130; DB 15; Length 114;  
 Best Local Similarity 75.0%; Pred. No. 1.5e-12;  
 Matches 21; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ETLMOOQRLNSWGCKGRICVTSYKWN 30  
 DB 31 ETLMOOQRLNSWGCKGRICVTSYKWN 58

## RESULT 8

AC 091EC5 PRELIMINARY; PRT; 216 AA.  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE GP41 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF07;  
 RA - Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL EMBL: AJ236394: CAB96243.1; to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ236394: CAB96243.1; to the EMBL/GenBank/DBJ databases.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam: PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 216 216  
 SQ SEQUENCE 216 AA; 25027 MW; 413AE9BF1B4FC9A CRC64;

## Query Match

Best Local Similarity 74.6%; Score 129; DB 15; Length 216;  
 Matches 20; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ETLMOOQRLNSWGCKGRICVTSYKWN 30  
 DB 39 ETLIONOOLNSWGCKGRICVTSYKWN 66

## RESULT 9

ID 091EB6 PRELIMINARY; PRT; 219 AA.  
 AC 091EB6;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE GP41 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCF14;  
 RA Roques P., Robertson D., Sandrine S., Christel D., Francois S.,  
 RA Philippe M.;  
 RT "Phylogenetic analysis and sub-typing of 47 HIV-1 group O isolates."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ236404: CAB96252.1; to the EMBL/GenBank/DBJ databases.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam: PF00517; GP41; 1.  
 KW Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 219 219  
 SQ SEQUENCE 219 AA; 25353 MW; F65829EB01520D91 CRC64;

Query Match 74.6%; Score 129; DB 15; Length 219;  
 Best Local Similarity 71.4%; Pred. No. 4.4e-12;  
 Matches 20; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSGCKGRILCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 49 ETLIQNQLNLMSGCKGRILCYTSVKWN 76

## RESULT 10

O11940 PRELIMINARY; PRT; 120 AA.

AC 011940;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-ABR1123;  
 RX MEDLINE-97340911; PubMed-9197385;

RA Brennan C.A., Hackett J. Jr., Zekeng L., Lund J.K., Vallari A.S.,  
 Hickman R.K., Gurtler L., Kapue L., von Overbeck J., Hampl H.,  
 Devare S.G.;

RT "Sequence of gp120 immunodominant region of HIV type 1 group O from  
 west central Africa."

RL AIDS Res. Hum. Retroviruses 13:901-904(1997).

DR EMBL: U90133; AAB62816.1; -;  
 DR InterPro: IPR000328; Env\_Gp41.

DR Pfam: PF00517; GP41; 1.

KW Transmembrane.

FT NON\_TER 1 1  
 FT NON\_TER 120 120

SEQUENCE 120 AA; 14463 MW; 22068B4D28222E44 CRC64;

Query Match 74.0%; Score 128; DB 15; Length 120;  
 Best Local Similarity 75.0%; Pred. No. 3.2e-12;  
 Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSGCKGRILCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 30 ETLIQNQLNLMSGCKGRILCYTSVKWN 57

## RESULT 11

O91HU1 PRELIMINARY; PRT; 134 AA.

AC 091HU1;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN GP41.

OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-97CM747;  
 RX MEDLINE-20386754; PubMed-10933623;

RA Wang C., Gao F., Fontjuno P.N., Zekeng L., van der Groen G.,  
 Plehazek D., Schable C., Lal R.B.;

RT "Phylogenetic analysis of protease and transmembrane regions of HIV  
 type 1 group O."

RL AIDS Res. Hum. Retroviruses 16:1075-1081(2000).

DR InterPro: IPR000328; Env\_Gp41.  
 DR Pfam: PF00517; GP41; 1.

KW Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 134 134

SEQUENCE 134 AA; 16066 MW; F6CBF09ABBF87403 CRC64;

Query Match 74.0%; Score 128; DB 15; Length 134;  
 Best Local Similarity 75.0%; Pred. No. 3.7e-12;  
 Matches 21; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSGCKGRILCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 32 ETLIQNQLNLMSGCKGRILCYTSVKWN 59

## RESULT 12

O9WRV5 PRELIMINARY; PRT; 143 AA.

AC 09WRV5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-ESP1;  
 RX MEDLINE-99284364; PubMed-10357472;

RA Mas A., Quiñones-Mateu M.E., Domingo E., Soriano V.;

RT "Phylogeny of HIV type 1 group O isolates based on env gene  
 sequences."

RL AIDS Res. Hum. Retroviruses 15:769-773(1999).

DR EMBL: AF081814; AAD41226.1; -;  
 DR InterPro: IPR000328; Env\_Gp41.

DR Pfam: PF00517; GP41; 1.

KW Transmembrane.

FT NON\_TER 1 1  
 FT NON\_TER 143 143

SEQUENCE 143 AA; 16325 MW; 1E44A4C631BA10A2 CRC64;

Query Match 74.0%; Score 128; DB 15; Length 143;  
 Best Local Similarity 75.0%; Pred. No. 3.9e-12;  
 Matches 21; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 ETLMDQORLNSGCKGRILCYTSARWH 30  
 |||:|||||:|||||:|||||:|||||:|:  
 Db 61 ETLIQNQLNLMSGCKGRILCYTSVKWN 88

## RESULT 13

O9WRV4 PRELIMINARY; PRT; 143 AA.

AC 09WRV4;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ENVELOPE GLYCOPROTEIN (FRAGMENT).

GN ENV.

OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-ESP2;  
 RX MEDLINE-99284364; PubMed-10357472;

RA Mas A., Quiñones-Mateu M.E., Domingo E., Soriano V.;

RT "Phylogeny of HIV type 1 group O isolates based on env gene  
 sequences."

RL AIDS Res. Hum. Retroviruses 15:769-773(1999).

DR EMBL: AF081815; AAD41227.1; -;  
 DR InterPro: IPR000328; Env\_Gp41.  
 DR Pfam: PF00517; GP41; 1.

Search completed: June 20, 2002, 15:34:36  
Job time: 214 sec

Search completed: June 20, 2002, 15:34:36  
Job time: 214 sec

Fri Jun 21 10:39:08 2002

us-09-605-573a-69.rspt